AD_____

Award Number: DAMD17-00-1-0204

TITLE: p53 Gene Mutagenesis in Breast Cancer

PRINCIPAL INVESTIGATOR: Steve S. Sommer, M.D., Ph.D.

CONTRACTING ORGANIZATION: Beckman Research Institute of
City of Hope
Duarte, CA 91010

REPORT DATE: March 2005

TYPE OF REPORT: Final

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

20050603 222

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503

1. AGENCY USE ONLY	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	March 2005	Final (1 Aug 2000 - 1 Feb 2004)

4. TITLE AND SUBTITLE
p53 Gene Mutagenesis in Breast Cancer

5. FUNDING NUMBERS
DAMD17-00-1-0204

6. AUTHOR(S)

Steve S. Sommer, M.D., Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Beckman Research Institute of
City of Hope
Duarte, CA 91010

E-Mail: sommeradmin@coh.org

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 Words)

The incidence and mortality from breast cancer varies significantly in ethnically and geographically distinct populations from the United States, Western Europe and Japan. **The central hypothesis of this proposal is that variability in the patterns of p53 mutagenesis in breast cancer reflects differences in exposures to different amounts and/or types of diverse environmental mutagens.** We postulate that mammary cells are exquisitely sensitive to lipophilic mutagens in the diet because of the unique architecture of breast tissue, i.e., tiny islands of cancer-prone rapidly dividing mammary cells surrounded by a sea of fat cells. It has now become possible to measure mutation load in an individual by identifying p53 mutations in the **normal** mammary cells in an individual. This new tool greatly increases the power to test the central hypothesis. We propose to test twenty upper Midwest U.S. women with breast cancer, ten with diets high in animal fat and ten with diets low in animal fat. In each woman, 30 different mutations will be defined in mammary cells. The normal breast tissue is obtained away from the tumor margin in women with breast cancer. With this newly developed technique, it is possible to determine the mutation fingerprint in each woman and correlate this with her diet type. It is hypothesized that there will be substantial individual variation in mutation patterns within the population consistent with a key prediction of the "lipophilic mutagens hypothesis".

14. SUBJECT TERMS	15. NUMBER OF PAGES
Molecular epidemiology, cancer etiology, mutation analysis, breast cancer	66
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

Table of Contents

INTRODUCTION

The incidence and mortality from breast cancer varies significantly in ethnically and geographically distinct populations (Boring et al., 1994; Ziegler et al., 1993). Patterns of p53 gene mutations in breast cancers are different among several ethnically and/or geographically distinct populations from the United States, Western Europe and Japan (Blaszyk et al., 1996; Hartmann et al., 1997). **The central hypothesis of this proposal is that variability in the patterns of p53 mutagenesis in breast cancer reflects differences in exposures to different amounts and/or types of diverse environmental mutagens (Biggs et al., 1993; Blaszyk et al., 1996).** As stated in our application, we strongly suspect that these toxins would be derived from the diet (Blaszyk et al., 1996; Hartmann et al., 1997). We planned to test one prediction of this hypothesis: that the p53 mutation pattern in breast cancers in the 1960's would have a different pattern of mutation than those previously defined by us in the 1990's.

Mutational patterns may be expected to change as new and/or a greater amount of mutagens are introduced into the environment and as other mutagens are eliminated. In particular, there have been major dietary changes in the past 30 years in this population and this may be expected to change the pattern of mutagen exposure. Indeed, we postulated that mammary cells are exquisitely sensitive to lipophilic mutagens in the diet because of the unique architecture of breast tissue, i.e., tiny islands of cancer-prone rapidly dividing mammary cells surrounded by a sea of fat cells. A limitation of the experiments that were initiated is that a population as a whole needed to be analyzed for some average effect in the population, since at most one mutation could be defined per individual. It has now become possible to measure mutation load in an individual by identifying p53 mutations in the **normal** mammary cells in an individual.

This new tool greatly increases the power to test the central hypothesis. We propose to test twenty upper Midwest U.S. women with breast cancer, ten with diets high in animal fat and ten with diets low in animal fat. In each woman, 30 different mutations will be defined in mammary cells. The normal breast tissue is obtained away from the tumor margin in women with breast cancer. With this newly developed technique, it is possible to determine the mutation fingerprint in each woman and correlate this with her diet type. It is hypothesized that there will be

substantial individual variation in mutation patterns within the population consistent with a key prediction of the "lipophilic mutagens hypothesis".

1. Terminology

Mutation load – the frequency, pattern and spectrum of somatic mutations

SCIMLA – Single Cell Immunohistochemical Mutation Load Assay. The analysis of mutation load, pattern and spectrum using single cells stained and microdissected from human tissue

Red cell – a cell with nuclear immunohistochemical staining for p53 overabundance

Clear cell – a cell without nuclear immunohistochemical staining for p53 overabundance

Sequence change – any DNA base change

Mutation frequency – the number of cells with a unique conserved, somatic missense change in the total number of cells examined

Mutant frequency- the number of red cells estimated by the number of red cells divided by the total number of cells screened.

Mutation pattern – the relative frequency of different types of somatic mutations

Mutation spectrum – the somatic mutation frequency and type at each base position

ADO – allele drop out; defined as the failure to detect one of the two autosomal alleles.

Pyrophosphorolysis: removal of the 3' terminal nucleotide by DNA polymerase in the presence of pyrophosphate (PPi) to generate the nucleotide triphosphate. This is the reverse of DNA polymerization. Pyrophosphorolysis is the reverse reaction of DNA polymerization. In the presence of pyrophosphate, the 3' nucleotide is removed from duplex DNA to generate a triphosphate nucleotide and a 3' terminal shortened duplex DNA: $[dNMP]_n + PPi \rightarrow [dNMP]_{n-1} + dNTP$(Duetcher and Kornberg 1969).

PAP: Pyrophosphorolysis activated polymerization.

P*: an oligonucleotide with a blocked 3' terminus that is not directly extendable but is activatable by pyrophosphorolysis.

Sensitivity: The sensitivity is defined as the minimum copy number of a template that generates a detectable product when P*s match the template. It is determined by serial dilution of the matched template.

Specificity: The specificity is defined as the maximum copy number of the mismatched template that generates an undetectable product when P*s mismatch the template. It is determined by serial dilution of the mismatched template.

Selectivity: The selectivity is the ratio of sensitivity to specificity.

SCIMPLA – Single Cell Immunohistochemical Partial Mutation Load Assay Simplified version of SCIMLA in which mutant load is determined in multiple animals (e.g. 20) in which average mutation load is estimated from distribution of mutant loads. SCIMPLA does not require microdissection or sequencing.

2. Cancer is a Disease of Mutations

Most, if not all mutagens are carcinogens, although there are carcinogens that are not obviously mutagens. A better tool for measuring the consequences of mutagens in the environment or endogenous defects is needed. There are available systems for determining the underlying pattern of mutation in humans, mouse and fish. When spontaneous mutation is compared in mouse and Medaka fish in the same transgenic system, the mutation frequency pattern and spectrum is remarkably similar as it is to the deduced underline pattern of spontaneous human mutations, as deduced by analyses of human Factor IX gene. However, for most human genes defective in inborn errors of metabolism, the biology of the gene and the biology of the disease often skew the underlying pattern so that the observed pattern in disease can vary markedly.

3. Mutations in the *p53* gene

Signature mutations permit identification of mutagens. The mutational process can be characterized by the resulting pattern of mutation, defined as the relative frequencies of various categories of mutation (e.g., transitions, transversions, deletions and insertions). Virtually all studied chemical mutagens alter the mutation pattern in specific ways.(Friedberg et al. 1995) Spontaneous mutation patterns in transgenic and endogenous targets are distinct from mutation patterns following exposures to known mutagens. For example, benzo[a]pyrene (B[a]P) and ethylnitrosourea (ENU) specifically cause point mutations at G:C and A:T base pairs, respectively.(Skopek et al. 1996; Walker et al. 1996) Dimethylbenzanthracene (DMBA) causes primarily base pair substitutions, predominantly A:T to T:A (44%) and G:C to T:A (24%).(Manjanatha et al. 2000) In human populations, signature mutations associated with

specific exogenous mutagens have been observed in samples of somatic mutations in the *p53* tumor suppressor gene(Harris and Hollstein 1993; Hussain and Harris 1999), e.g., G to T transversions in lung cancers of cigarette smokers(Chiba et al. 1990), C to T and CC to TT tandem dipyrimidine transitions in skin cancers associated with excessive sun exposure(Brash et al. 1991) and G to T transversions at arginine 249 in liver cancers associated with aflatoxin exposure.(Bressac et al. 1991; Hsu et al. 1991; Scorsone et al. 1992; Aguilar et al. 1993; Li et al. 1993).

Such signature mutations can be detected in the mutation pattern with high sensitivity. Patterns of mutations can vary dramatically. As illustration, in a comparison of the pattern of *p53* mutations in lung cancers of cigarette smokers with that in colon cancer and the germ line, a highly significant difference ($P < 0.0001$) could be detected even though only nine mutations in lung cancers were available.(Sommer 1990) In the previous examples, one specific mutagen is believed to have had a dominant influence on the mutation pattern. In general, differences in the mutation patterns between populations are due to differences in both endogenous processes and exogenous mutagen exposures, and interactions between these influences. Furthermore, mutagen exposure may generally involve a mix of mutagens with comparable degrees of influence on the mutation pattern, rather than one dominant mutagen. Differences between the mutation patterns in population samples resulting from such complex differences can be detected.

4. The *p53* gene is a good mutation reporter for a variety of reasons

Alterations of the *p53* gene are the most frequent genetic abnormalities in human malignancies.(Greenblatt et al. 1994) The frequency of *p53* mutations in various cancer tissues ranges from 10% to 80%.(Soussi et al. 1994) The precise mechanisms by which *p53* protein carries out its biological function are not clear. The ability of *p53* to bind to specific DNA sequences and to activate transcription suggests an important role in the regulation of cell proliferation.(Harris 1996a; Harris 1996b) Loss of *p53* function eliminates the growth arrest response to DNA damage(Kastan et al. 1992) and enhances the frequency of gene amplification,(Livingstone et al. 1992; Yin et al. 1992) suggesting a role for *p53* in the control of a cell-cycle checkpoint and maintenance of the integrity of the genome.(Lane 1992) *p53*

controls the G1 cell cycle checkpoint by activating transcription of genes with *p53*-responsive elements.

The products of these genes trigger G1 arrest after DNA damage, preventing replication of a damaged DNA template. *P53* protein is hypothesized to function in the maintenance of genomic stability by blocking cells from replication after DNA damage until the damage is repaired or if the damage is extensive by initiating apoptosis.(Lane 1992) *P53* has several key advantages as a model mutation reporter. 1) Mutations in *p53* are clustered within exons 5 to 8 (80% of mutations within exons 5 to 8;(Hartmann et al. 1995a). 2) *P53* protein is stabilized by missense mutations primarily occurring in exons 5 to 9. *P53* protein stabilization permits immunohistochemical staining to identify cells with mutant *p53*.(Bartek et al. 1990; Prosser et al. 1990; Bartek et al. 1991; Davidoff et al. 1991; Varley et al. 1991; Bartek et al. 1993) 3) Mutation screening and DNA sequence determination of exons 5 to 9 involves reasonable effort.(Kovach et al. 1991; Blaszyk et al. 1995) 4) A great deal of information on the degree of sequence conservation is available.(Soussi et al. 1990; Walker et al. 1999) Despite speculations that the *p53* gene is "the guardian of the genome",(Lane 1992) compelling data indicate that the frequency and pattern of mutation are generally unchanged in normal and tumor tissues of *p53* nullizygous mice that develop cancer early.(Sands et al. 1995; Buettner et al. 1996; Buettner et al. 1997) This evidence suggests that *p53* is a good mutation reporter because the background rate of mutations in *p53* deficient cells is similar to that of wild type cells.

5. Breast cancer incidence is high, disease etiology is unknown but epidemiological studies support a role for environmental mutagens

Breast cancer is the most frequent malignancy occurring in women worldwide, with some 600,000 new cases arising each year and accounts for nearly 20% of all cancers among women.(Higginson et al. 1992) Breast cancer is the major cause of cancer-related deaths in women residing in the Western world.(Coleman 2000; Kaplan and Wingard 2000) The worldwide incidence of breast cancer varies by at least fourfold.(Coleman et al. 1993; Coleman 2000) The incidence and mortality from breast cancer varies significantly in ethnically and geographically distinct populations. Northern peoples, including Scandinavians, Scotts, Welsh, English and North Americans, have the highest rates of breast cancer whereas Italians, Spaniards

and Central and South Americans have lower rates of breast cancer.(Boring et al. 1994) Striking evidence for the geographic specific factors associated with breast cancer incidence come from altered breast cancer incidence for emigrants from low-risk populations to high-risk population.(Stemmermann 1991; Coleman et al. 1993; Coleman 2000) Asians, in general, have had very low rates of breast cancer in the past, which increase upon immigration to the West.(Ziegler et al. 1993).

This effect is seen in the first generation and is more dramatic in second and subsequent generations of Japanese in the United States. Despite intensive study, the origins of sporadic breast cancer are largely unknown.(Lane 1992) Studies examining the role of specific exposures in breast cancer have found either inconsistent or weak associations.(Millikan et al. 1995) It would be predicted that epidemiological studies would be inconsistent if different mutagens predominate in different high-risk cohorts. Analysis of the pattern of mutation in an endogenous human gene generally should allow high-risk cohorts with different mutagen exposures to be distinguished from each other due to the different signatures of mutagens.

6. The *p63* and *p73* genes are used as mutation reporters

While *p53* is a single predominant protein of 393 amino acids coded by 11 exons, *p63* and *p73* have three extra exons, coding for an extra stretch of 250 residues. There is an amino acid identity of 65% between the DNA binding domains of *p53*, *p63* and *p73* proteins. *P63* and *p73* proteins are predominately located in the nucleus (Inoue et al. 2002). Mutant p73 protein is more stable than the wild type. (Inoue et al. 2002).

BODY

1.The Sommer laboratory has recently developed a prototypical *p53* mutation load assay

A flowchart of the methodology of the prototypical *p53* mutation load assay (SCIMLA) is presented in Figure 1. The *p53* gene is used as a model reporter because *p53* protein is stabilized and accumulated by missense mutations in exons 5 through 9, making possible immunohistochemical staining of mutant cells.(Bartek et al. 1990; Prosser et al. 1990; Bartek et al. 1991; Davidoff et al. 1991; Varley et al. 1991; Bartek et al. 1993). Ethanol-fixed, paraffin-embedded tissues were immunostained with *p53* antibodies (Heinmoller et al. 2002). Single

cells that stained positively for *p53* were microdissected and exons 5 to 9 of the *p53* gene were amplified from genomic DNA using a Stimulated-PCR method (Schlake et al. 2003). Tissue processing was optimized to maintain DNA integrity (Heinmoller et al. 2002). Single stranded genomic DNA was determined to have an average size of 20 kb enabling the amplification and analysis of the 1.6 kb region of the *p53* gene. Stimulated PCR has a success rate of 80 to 90% and a high yield with minimal primer artifacts. Until recently, only a few studies have demonstrated successful molecular analysis using single cells dissected from paraffin-embedded tissues. The results in these studies were limited by low amplification efficiency and high rates of allele drop out (ADO; reviewed in Heinmoller et al., Appendix).

Tissue fixation with ethanol (85%) and the addition of 0.2 mM EDTA helped to achieve 80 to 90% success in amplification of 1.6 kb from a single cell from paraffin-embedded tissue. Normal tissue sections were immunohistochemically stained for overabundance of *p53* protein. Microdissection of single cells was performed with a manual micromanipulator equipped with a Tungsten needle. Sequence analysis of the *p53* gene was performed following Stimulated PCR using single microdissected cells. These technical advances facilitate routine mutation analysis using a single cell microdissected from routinely processed paraffin-embedded normal and tumor tissues. Allele drop out still represents a serious problem in single cell mutation analysis, especially in samples with limited template DNA and prone to DNA-damage (Figure 3).

2. The prototypical SCIMLA identifies missense mutations in positively stained cells with a low rate of false positives

Initial results:

43 cells staining positive for *p53* overabundance have been manually microdissected and a 1.6 kb region of the *p53* gene amplified and sequenced to identify sequence changes (including exons 5 to 9; Table 1). A total of 21 sequence changes were observed in the 43 red cells analyzed and 15 of these changes were conserved missense changes (35%). A total of 34 cells not staining for *p53* overabundance were also manually microdissected and only 5 sequence changes and 1 conserved missense change (3%) were identified in the same 1.6 kb region of the *p53* gene. Both alleles were detected in 29 of 48 cells (60%) with a known polymorphism. Thus 40% ADO was measured for our novel approach (Figure 3). For the 34 clear cells analyzed, there were five sequence changes observed. These five changes, which are not in the germline, are each found in single cells and are putatively the result of polymerase error, although

there is a very small chance that any given cell might be a rare cell chosen at random that happened to have a rare somatic mutation in the *p53* gene but did not stain immunohistochemically. Thus we estimate the rate of polymerase error to be 0.00011. For the red cells, a second estimate of polymerase error can be made by focusing on all the nucleotides that would not give rise to missense changes at conserved amino acids (1359 nucleotides).

There were five changes in 43 cells that were not missense changes at conserved amino acids. The target size for amino acid for missense changes at evolutionary conserved residues is 310 nucleotides. Silent changes and intronic changes in the region sequenced total 1,359 nucleotides. Thus, the ratio of non-missense changes to missense change at evolutionary conserved amino acids is expected to be 4.4, similar to the five to one ratio found in the 34 clear cells. Each of the changes is in a different cell (14% of all cells) and we conclude that these are putative polymerase errors. Thus, there are five errors in 58,437 nucleotides (43 cells x 1,359 nucleotides) or 0.00010. Thus, two different estimates indicate that the rate of polymerase error is 0.0001.

3. To date a fully integrated SCIMLA is available for optimization and adaptation for high throughput.

In summary, the initial studies for SCIMLA define the fixative, section thickness, immunohistochemical staining protocol, microdissection technique, PCR conditions, post PCR purification and sequencing protocols optimized to date. If 35% of red cells contain a mutation and 43% of polymorphic cells contain allele drop out, it is anticipated that roughly 57% of the samples would have a mutation in the absence of allele drop out. The majority of the red cells are already expected to have a mutation.

4. Mutations load in normal breast tissue in 3 patients:

As part of our preliminary studied we attempt to measure the mutation load in normal breast tissues in three patients. The following tables summarize the results. (Table 3).

5. Immunohistochemical staining of *p63* and *p73* proteins

For *p73* immunohistochemical staining a rabbit polyclonal antibody *p73* (Figure 4) (Chemicon, Australia) was used at a 1:100 dilution to detect *p73* protein. *P63* protein was stained with a mouse anti-p63 monoclonal antibody (dilution 1:100, Chemicon, Australia). *P63* and *p73* antibodies were used to stain various tissues. DAB (3,3` -diaminobenzidine tetrahydrochloride), which develops the stain brown, was used as a chromogen. The cells were considered positive when a strong brown nuclear staining was detected. The other "clear" cells were negative. On each slide the number of epithelial cells ranged

between 1,000 and 25,000. The number of positive stained cells by *p63* and *p73* is about 1 per 100,000 cells under optimal conditions. In breast tissue and occasionally in colon and lung tissues we found a strong nuclear staining within a continuous layer of the basal ductal epithelium, single crypt cells, and bronchiolar epithelium, respectively, but these cells were not counted.

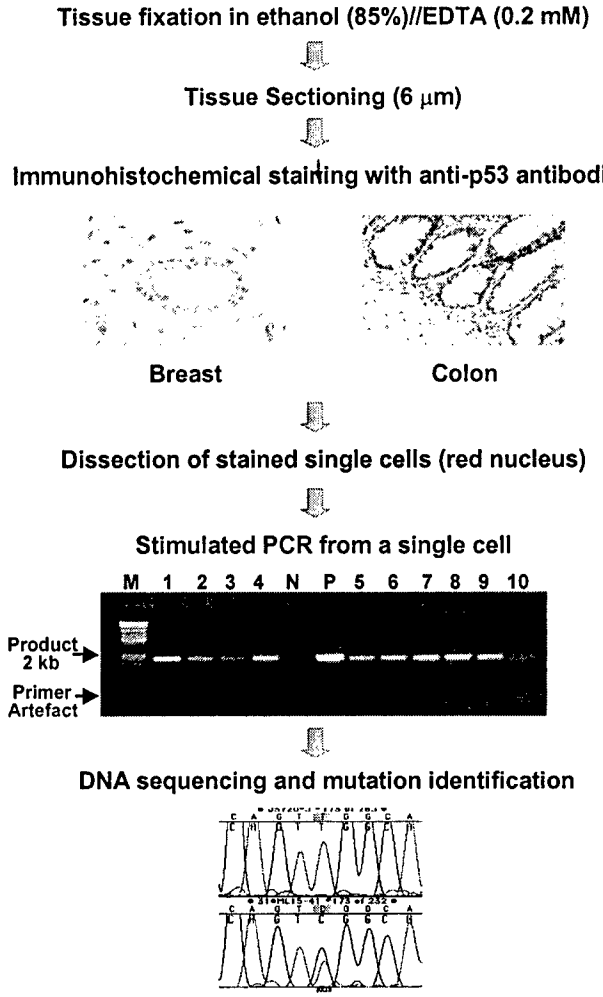


Figure 1. A flow chart of SCIMLA

I. Modified tissue processing preserves more intact nuclei and genomic DNA. II. Immunohistochemistry staining normal breast and colon tissue identifies mutant cells with frequency of 2.5×10^{-4}. *p53* positive staining is indicated in bright red color and no staining is in clear color. III. Microdissection of stained single cells selects more intact nuclei. IV Stimulated PCR from single cells is with 87% successful percentage. A 1.8 kb region of the *p53* gene was amplified from genomic DNA of microdissected single cells for 40 cycles. Lanes 1 to 10 are single cell samples. Lane N = No DNA control. Lane P = positive control with 20 cells, M = DNA marker. The PCR product and primer artifacts of 200 bp are indicated. V. Sequence analysis to detect mutations shows a typical pattern of the wild type and a heterozygous mutation from single cells. The C mutation is identified with the C peak area ≥20% of the wild type T peak.

Table 1. Somatic mutations detected by SIMLA

Individual ID[1]	Cell ID	Sequence Change[2]	Amino Acid Change	Species Conservation[3] monkey	mouse	rat	chicken	Xenopus	trout
	Red Cells								
1	1	G14617A	NC[4]						
1	2	A14582G	Lys=>Arg	Y	Y	Y	Y	Y	Y
2	3	A13419G	Tyr => Cys	Y	Y	Y	Y	Y	Y
2	4	G13238A	Gly => Ser	Y	Y	Y	Y	N	Y
2	5	A13419G	Tyr => Cys	Y	Y	Y	Y	Y	Y
2	6	A13245G	NC[4]						
2	6	T14675C	NC[4]						
2	7	A13419G	Tyr => Cys	Y	Y	Y	Y	Y	Y
2	8	A13419G	Tyr => Cys	Y	Y	Y	Y	Y	Y
2	9	A13419G	Tyr => Cys	Y	Y	Y	Y	Y	Y
2	10	A13419G	Tyr => Cys	Y	Y	Y	Y	Y	Y
2	11	A13419G	Tyr => Cys	Y	Y	Y	Y	Y	Y
2	12	C14793T	NC[4]						
2	13	A13419G	Tyr => Cys	Y	Y	Y	Y	Y	Y
2	13	A14067G	Glu => Gly	Y	Y	Y	Y	Y	Y
2	14	A13419G	Tyr => Cys	Y	Y	Y	Y	Y	Y
3	15	C13155T	Ala => Val	Y	Y	Y	Y	Y	Y
3	16	C13155T	Ala => Val	Y	Y	Y	Y	Y	Y
3	17	C13155T	Ala => Val	Y	Y	Y	Y	Y	Y
3	18	C13155T	Ala => Val	Y	Y	Y	Y	Y	Y
4	19	T13377G	Leu=>Trp	Y	N	Y	Y	N	N
	Clear cells								
1	1	G13637A	NC[4]						
2	2	C13984T	NC[4]						
3	3	C13253T	NC[4]						
3	4	T13413C	Val=>Ala	Y	Y	Y	Y	N	N
3	4	C14283T	NC[4]						

[1]Data are from normal breast tissue obtained beyond the tumor margin in four women who underwent surgery for removal of a primary breast tumor

[2]The genomic sequence of the p53 gene was from a revised version of X54156 in GenBank;
The sequence changes were not present in the germline of these individuals as determined by sequencing of DNA isolated from 1,000 cells

[3]Species conservation was determined according to Soussi et al. [1990]; a conserved missense change is defined as identity in monkey, mouse, rat and chicken

[4]NC = no known functional or structural change

Table 2. Comparison of unique mutation pattern in three SCIMLA patients.

Patient	Pattern Comparisons by Fisher's Exact Test				Transisitions at CpG	Transitions not at CpG		Transversions			
	ABC 3x7	AB 2x7	BC 2x7	AC 2x7	G:C> A:T	G:C> A:T	A:T> G:C	G:C> C:G	G:C> T:A	A:T> C:G	A:T> T:A
A	0.046*	0.29	0.13		6	3	0	0	1	0	1
B					6	2	3	0	4	0	0
C				0.017*	6	0	5	0	0	0	0

Tissues previously analyzed by the extensively validated Big Blue® system will be used to help validate SCIMLA. The laboratory has more than 10 years of experience in measuring and analyzing mutation load with the Big Blue® system. We have measured mutation load in multiple tissues through the life cycle of the mouse. The effects of p53 nullizygosity have been analyzed (Buettner et al. 1996). Analyses of tandem, multiple mutations and microdeletions/microinsertions have been of particular interest. Big Blue references: (Hill et al. 2003; Hill et al. 2004a; Hill et al. 2004b). SCIMLA has the advantage of analyzing endogenous gene in a diversity of animals, including animals taken from the wild. With SCIMLA, mutation load can be analyzed in any Knock out/knock in mouse without breeding the *Lac I* transgenic mouse. Fig. 5 shows a comparison of mutant frequency with mutation frequency in two tissues (testis and cerebellum), in Big Blue® mice.

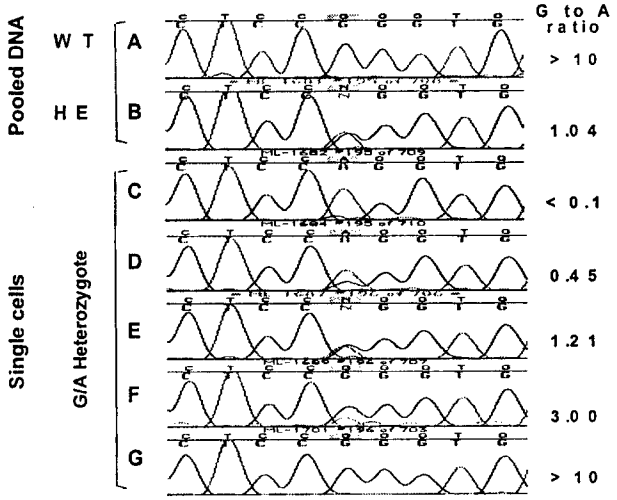


Figure 2. Peak patterns in a known heterozygous G to A polymorphism

Single cells containing a known heterozygous G to A polymorphism in exon 5 at position 13494 were taken as an example. The amplified product was sequenced using *p53*(13016)24D primer. Panel A is the wild type populated DNA of more than 1000 copies. Panel A polymorphism amplified from□B is the positive control of the heterozygous G A□populated DNA. Panels C to G are the different patterns of the heterozygous G polymorphism amplified from single cells. The ratio of G to A peak area is indicated. The peaks areas of the G and A signals are varied with single cell samples from <0.10 to >10 in which either G or A is lost.

Fig 3A

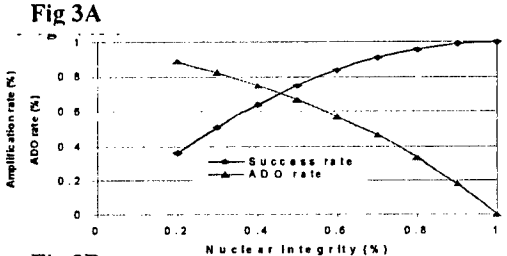


Fig 3B

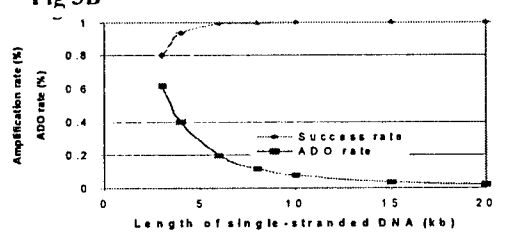


Figure 3 Effect of fraction of nucleus volume (A) and single-stranded DNA length (B)

A. The effect of fraction of nucleus volume on the amplification percentage and allele dropout rate. The fraction of nucleus volume (%), X-axis, is the cubic portion of nuclei from microdissection. The amplification percentage (%), Y-axis, is the ratio of the number of amplified cells to the total cells. The allele dropout rate (%), Y-axis, is the ratio of the number of allele dropout cells to the number of amplified cells. For example, the fraction of nucleus volume (%) is 70%, the amplification percentage is 91.0% and the allele dropout rate is 46.2%. **B. The length effect of single-stranded DNA on the amplification percentages and allele dropout rate.** X-axis is the average length of single-stranded DNA (kb). It is assumed that i) each single strand is of uniform size and nicks occur at random in each strand, ii) each single strand without nick is amplified equally. For example, if the length is 6 kb, a 2 kb PCR segment has 1/2 chance within an intact single-stranded template without nicking. The amplification percentage is 93.7% and the allele dropout rate is 40.0%.

Table 3. Analysis of mammary cells in three patients

	patient A	patient B	Patient C
# Mammary cells analyzed	1205000	1394400	912000
Number of red cells	94	37	20
Amplified	94	29	16
Discarded due to contamination	22	0	0
net amplified red cells	52	29	16
Non amplified	20	8	4
number of missense mutation	46	18	11
unique missense mutation	11	15	11
Mutant frequency	1/12819	1/38481	1/45600
mutation frequency	1/109000	1/92960	1/82909
Red cell amplification rate	73%	74%	80%
Percent of amplified red cells with missense mutations	88%	62%	68%
Percent of amplified red cell with Unique missense mutations	21%	52%	68%
Percent of silent mutation	0%	0%	0%

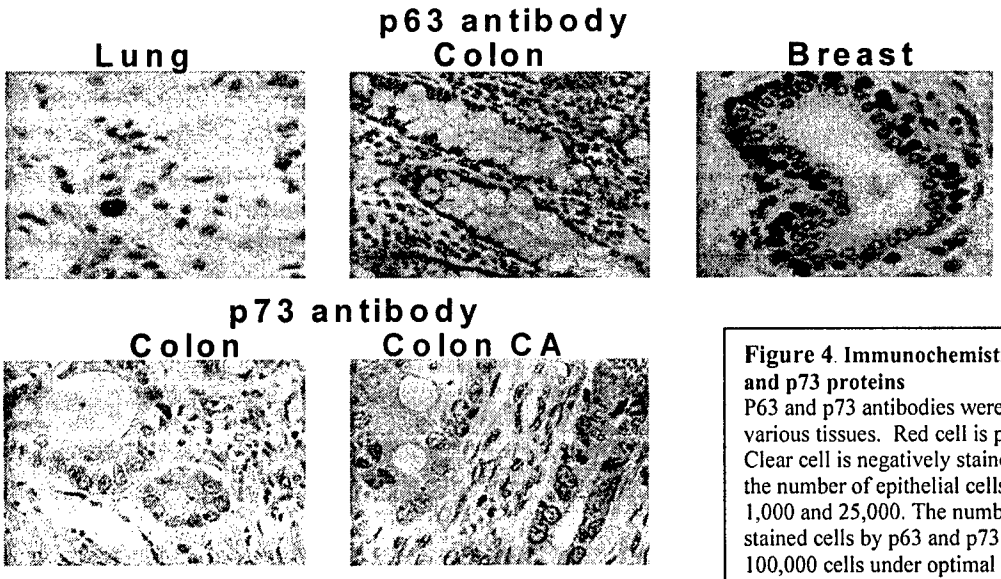


Figure 4. Immunochemistry staining of p63 and p73 proteins
P63 and p73 antibodies were used to stain various tissues. Red cell is positively stained. Clear cell is negatively stained. On each slide the number of epithelial cells ranged between 1,000 and 25,000. The number of positive stained cells by p63 and p73 is about 1 per 100,000 cells under optimal conditions.

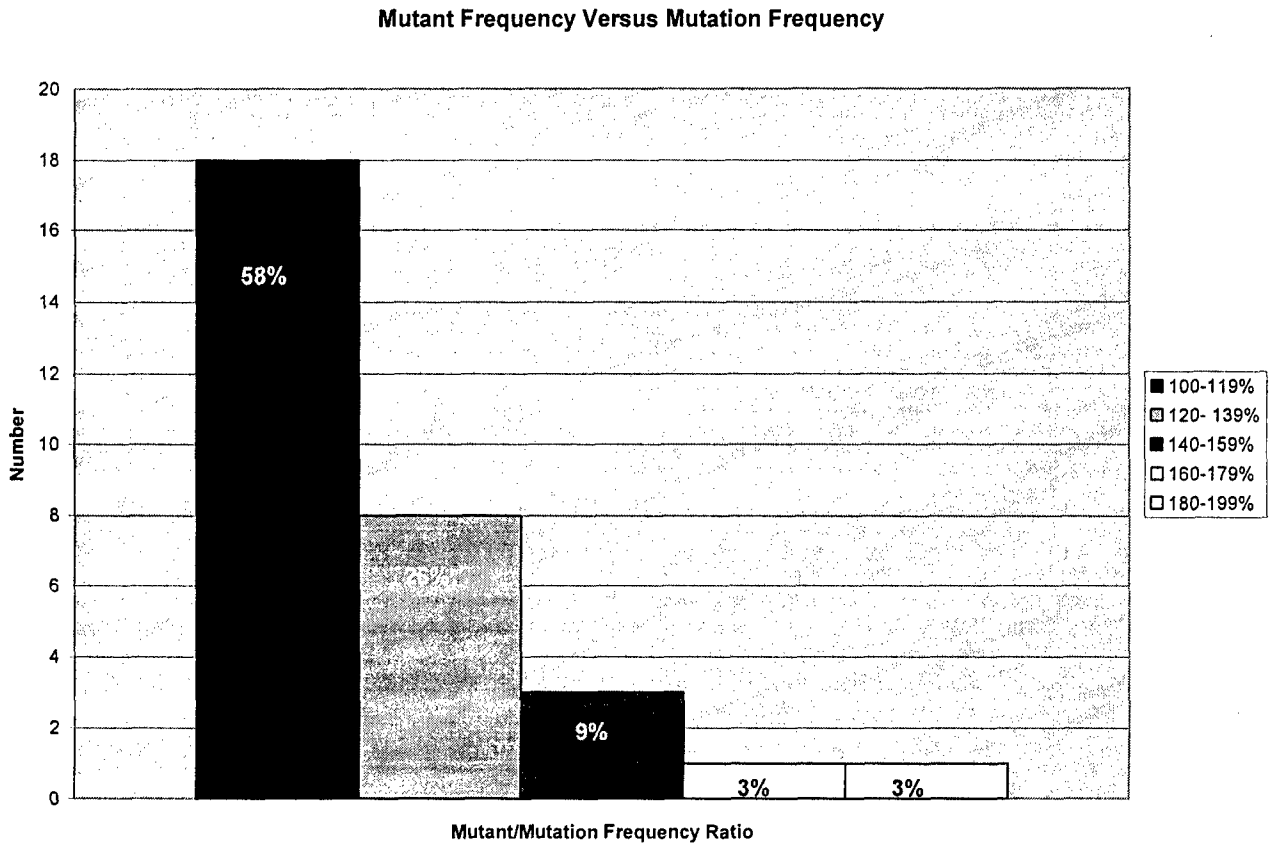


Fig 5. Mutant Frequency relative to Mutation Frequency in Big Blue® mice (testis and cerebellum combined).

6. PAP technology for detection of a single copy of rare mutations

Pyrophosphorolysis is the reverse reaction of DNA polymerization. In the presence of pyrophosphate, the 3' nucleotide is removed from duplex DNA to generate a triphosphate nucleotide and a 3' terminal shortened duplex DNA: $[dNMP]_n + PPi \rightarrow [dNMP]_{n-1} + dNTP$ (Duetcher and Kornberg 1969). PAP is a novel method for nucleic acid amplification (Liu and Sommer 2000) (Liu and Sommer 2002), where pyrophosphorolysis and polymerization are serially coupled by DNA polymerase using pyrophosphorolysis *activatable* oligonucleotides (P*). P*s are blocked at their 3' termini with dideoxy nucleotides that must be removed by pyrophosphorolysis for extension to occur. When P* is annealed to its complementary template strand, the 3' terminal blocker can be removed by pyrophosphorolysis in the presence of pyrophosphate. The activated oligonucleotide can then be extended by DNA polymerization.

The extreme selectivity of PAP results from the use of pyrophosphorolysis *activatable* oligonucleotides (P*) to serially couple pyrophosphorolysis and polymerization (Liu and Sommer 2000). Significant nonspecific amplification (Type II error) requires mismatch pyrophosphorolysis followed by misincorporation by the DNA polymerase, an event with a frequency estimated to be $3.3x10^{-11}$ (Liu and Sommer 2000). Great specificity of the primers without primer dimers may facilitate multiplex PCR in single cells (Fig 6). Bi-PAPA, an extension of PAP, can be used to measure ultra rare mutations for minimal residual disease detection and for mutation load analysis. Bi-PAPA complements SCIMLA; it is especially valuable for hotspots of mutation any nucleotide in the genome can be assayed. However, each nucleotide change requires a Bi-PAPA assay. Where as the target of mutations for SCIMLA, as per the available p53 databases includes at least 500 nucleotide changes.

7. <u>DOVAM-S for unknown mutation scanning (Fig 7)</u>

DOVAM-S, is a modification of SSCP in which up to 23 PCR products from genomic DNA are analyzed in a single lane under five different non-denaturing electrophoresis conditions, which include various matrices, buffers, temperatures, and additives (Liu et al. 1999a; Buzin et al. 2000). DOVAM-S analysis has been used to scan blinded analyses for *F9* and ATM and prospective analysis of *F8* genes with different sizes and numbers of exons. The redundancy of the multiple electrophoresis conditions permits the detection of virtually all mutations within the region scanned. The strongest advantages of the DOVAM-S method include the parallel analysis of 45 to 50 samples using multiple amplified products per lane each gel and the increase in sensitivity of mutation detection to approximately 100%.(Liu et al. 1999a). DOVAM-S is a robust method for scanning genomic DNA, particularly in large genes with widely separated exons and mutations scattered throughout the gene. This highly sensitive scanning technique has been used for mutation analysis in the hemophilia A, B and the DMD populations. We extensively validated DOVAM-S by blinded analyses estimate that about 98% of these mutations could be detected with three of the five conditions and 99.8% with all five conditions.

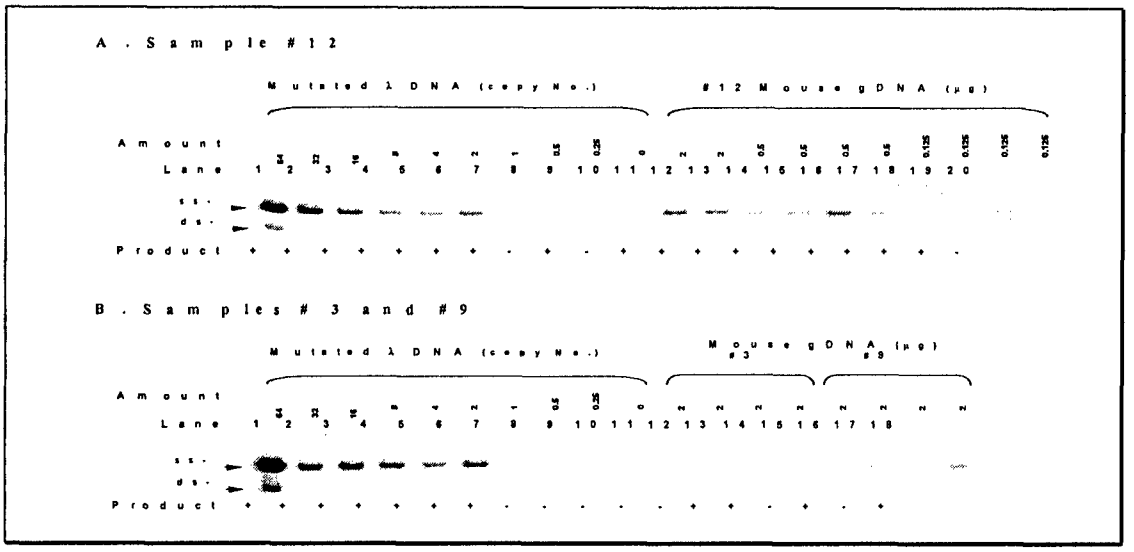
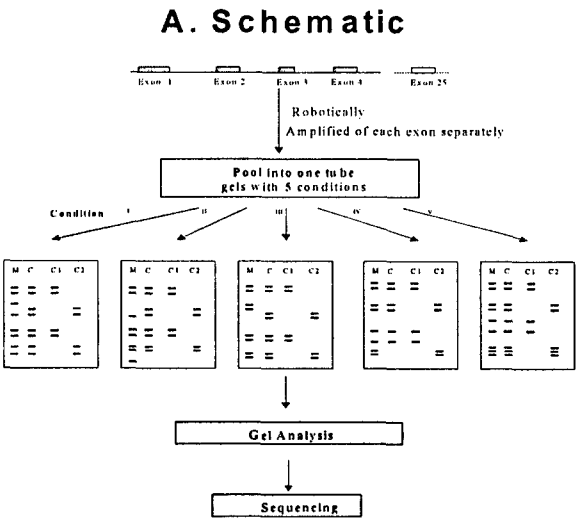


Figure 6. Detection of single molecule of target
Genomic DNA samples of the *lacI*$^+$ transgenic mice were chosen. 2 μg of mouse genomic DNA was amplified. A. Sample 12. Assay A was performed. In Lanes 11-12, 13-16 and 17-20, 2 μg, 0.5 μg and 0.125 μg of the *lacI*$^+$ mouse genomic DNA of sample 12 were used in each reaction, respectively. Lanes 1-10 are controls, the copy number of the mutated λ DNA per reaction was reconstructed by 2-fold serial dilutions. In lanes 1-10 and 13-20, each reaction also contained 1 μg of the *lacI*$^-$ mouse genomic DNA carrier. ss = single-stranded, ds = double-stranded.
B. Samples 3 and 9. Assay B was performed. In Lanes 11-14, 2 μg of the *lacI*$^+$ mouse genomic DNA of sample 3 was used in each reaction. In Lanes 15-18, 2 μg of the *lacI*$^+$ mouse genomic DNA of sample 9 was used in each reaction. Lanes 1-10 are controls, the copy number of the mutated λ DNA per reaction is indicated. Each control reaction also contained 1 μg of the *lacI*$^-$ mouse genomic DNA carrier.



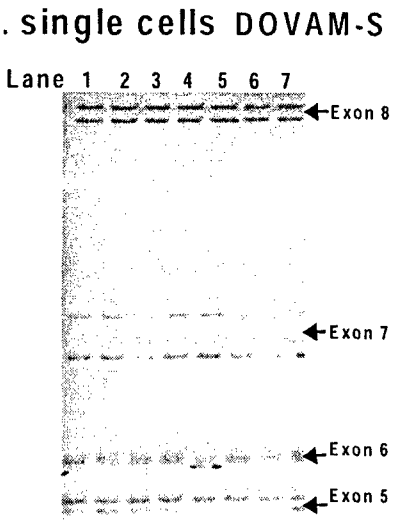


Figure 7. DOVAM-S detection of known mutations from single cells

A key advance has been the successful publication of manuscript stemming directly from the tasks of this proposal (Single-cell immunohistochemical mutation load assay (SCIMLA) using human paraffin-embedded tissues. Environ Mol Mutagen. 2003;42(3):206-15. Schlake G, Liu Q, Heinmoller E, Hill KA, Weiss L, Sommer SS.).

Key research advances i) improvement of the recent immunohistochemical protocol; ii) demonstration that an individual can harbor a large number of missense mutations in normal tissue a key point in the mutation load concept; and iii) Identification of a dietary questionnaire. The "Food and nutrition questionnaire of the breast cancer comprehensive project" was prepared originally for the national Action plan on breast cancer of the Office on Women's Health, U.S. Department of Health and human Services by Lenore Kohlmeier, Ph.D. from School of Public Health and human Services, University of North Carolina and by the institute of Survey research, Temple University. The questionnaire has the advantage of covering the dietary history over the lifetime of the subject rather than at a single period of time.

Task 1a. Development of novel methodologies to permit measurement of mutation load using the p53 gene in single cells from ethanol-fixed, paraffin-embedded and immunohistochemically-stained human breast tissue (months 1 to 12). Completed. Analysis of somatic mutation within single cells permits examination of an individual's mutation load, i.e., the overall mutation frequency and alterations in mutation pattern and spectrum. These measurements in normal tissues can identify individuals with a high mutation load at increased risk for cancer. Mutation frequency and type can vary dramatically with the type of mutagenic insult or defective repair system, and can provide clues to identify sources and or mechanisms of mutation.

The p53 gene was used as a model system, because the p53 protein is stabilized and accumulated by missense mutation in exons 5 through 9, making possible immunohistochemical staining of mutant cells. Ethanol-fixed, paraffin-embedded tissues were immunostained with p53 antibodies.

Single cells that stained positively for p53 were microdissected and exons 5 to 9 of the p53 gene were amplified from genomic DNA using a novel Stimulated-PCR method. Figure 1 is a flowchart outlining the novel approach to the measurement of individual mutation load. Stimulated PCR has a success rate of 80 to 90% and a high yield with minimal primer **artifacts.**

Task 1b. Measurement of mutation load in cells immunohistochemically stained for p53 overabundance and in cells not staining for p53 overabundance (months 1 to 12). Completed (Table 1). To date, 40 cells staining positive for p53 overabundance have been manually microdissected and a 2 kb region of the human p53 gene amplified and sequenced (including exons 5 to 9). A total of 24 sequence changes were observed in the 40 cells analyzed and 19 of these changes were conserved missense changes (47.5%). A total of 55 cells not staining for p53 overabundance were also manually microdissected and only 10 sequence changes and 3 (6%) conserved missense changes were identified in the same 2 kb region of the p53 gene.

Task 1c. Measurements of allele drop out (ADO) (months 1 to 12). Completed (Table 1). Both alleles were detected in 49 of 72 cells (68%) with a known polymorphism. Thus 32% ADO was measured for our novel approach to measure mutation load using the p53 gene in single cells microdissected from ethanol-fixed, paraffin-embedded and immunohistochemically-stained normal human breast tissue.

Task 2. We propose to test twenty upper Midwest U.S. women who are breast cancer patients, ten with diets high in animal fat and ten with diets low in animal fat. In each woman, 30 independent mutations will be identified in mammary cells of normal breast tissue. The mutation patterns for these two groups of women will be compared and possible mutagen fingerprints identified.

Research progress toward task 1

Development of the P53 Mutation Load Assay (see appendix for publication)

P53 Mutation Load Assay Materials and Methods

Immunohistochemical staining and microdissection

Fresh normal tissue from breast cancer patients was cut into 3 to 4 mm thin slices and immediately transferred into jars with the ethanol-based fixative (85 to 95% Ethanol with 0.2 mM EDTA buffer, pH 8.0) for at least 12 hours. The specimens were processed the following day, and paraffin embedded using standard procedures of routine pathology. From each tissue block, sections of 6-micron thickness were cut. The deparaffinization process included one Xylene step at room temperature for 30 minutes with shaking every 5 minutes, followed by steps in alcohols of different concentrations. During the Xylene step, the slides were every 5 minutes carefully shaken to intensify the deparaffinization process followed by hydration using washes in 100% to 70% alcohol. Steam heating at 96 to 100°C for 5 minutes in 1mM EDTA solution (pH 8.0) was performed to unmask the antigenic sites.

The p53 antibody (mouse monoclonal antibody DO-7, Novocastra) was used at a concentration of 1:150. **Another P53 antibody, PAB 240, was used in conjunction with the DO-7** to increase sensitivity. The tissue sections were stained immunohistochemically for nuclear overabundance of p53 protein. Cells showed either positive nuclear staining for p53 overabundance (bright red) or no staining (clear). Background fluorescent staining was very low. The counterstain was Hematoxylin. Single cells were microdissected manually using an inverted Microscope (Nikon TMS) and a mechanical micromanipulation system (Model MP-285 from Sutter Instruments). Under direct visualization of the immunohistochemical staining, a tungsten needle was manipulated using a joystick to separate the single cell from surrounding tissue. The positively stained single cell with intact nucleus was then picked up manually with a 271/2'' G needle, and transferred into a 0.2 ml PCR tube containing 5 :l digestion buffer (#3 High Fidelity Buffer without Mg^{++}, 2 mg/ml Proteinase K (Quiagen), 3% Tween-20 detergent and 0.2 mM EDTA, pH 8.0). For each new cell microdissected, a fresh needle was used in order

to avoid contamination. The single cell was digested at 50 °C for 16 hours with following inactivation of the enzyme at 90°C for 10 min. The remaining cells on the slide were scratched off, digested by Proteinase K and the genomic DNA was denatured and electrophoresed through a 1% agarose gel to determine the average length of single stranded DNA (i.e., a measure of DNA integrity). Genomic DNA quality was estimated by examination of the size of the single stranded DNA.

Stimulated PCR

Primer Design

Designing the primers is a critical step in a successful Stimulated-PCR reaction. All primers were designed and analyzed with Oligo5 software (National Biosciences). The T_m of the primer was estimated by the nearest neighbor method at 50 mM KCl and 250 pM DNA and T_m of the PCR segment was estimated by the formula of Wetmur: $T_{m\ product} = 81.5 + 16.6 \log (K+=0.05\ M) + 0.41\ (\%G + \%C) - 675/length$. The criteria for specificity included high-specificity with low base-pairing stability at the 3'end, no primer-dimer or hairpin formation more than 7 bases at the 3'end for any strand and any segment. The primer also had no false priming site on the mouse p53 gene to generate spurious products.

Stimulated-PCR conditions

Unless stated, the PCR mixtures contained a total volume of 25 :l: human genomic DNA from a microdissected single cell; #3 Expanded High Fidelity Buffer (Boehringer Mannheim); 3.5 mM MgCl2; 500 :M of each dNTP; 2% DMSO; 0.2 to 0.6 :M of each of the primers; a mixture of 1.25 U of Platinum *Taq* DNA polymerases High Fidelity (*Taq/GB-D*)/1.25U of Platinum *Taq* DNA polymerase (GIBCO BRL); 5 :g of BSA and 25 ng of mouse genomic DNA with the average size of more than 20 kb. The cycling conditions included denaturation at 92°C for 12 seconds, annealing at 60°C for 20 seconds, and elongation at 68°C for 2 minutes for 40 or 45 cycles with a Perkin Elmer GeneAmp PCR 9700 system. An additional 20 seconds of denaturation time proceeded the first cycle. PCR product (2 to 4 µl) was electrophoresed through a standard 1% agarose gel stained with ethidium bromide for UV photography using a CCD camera (Bio-Rad GelDoc 1000) and Multi-Analyst software (version 1.1).

PCR product purification, secondary amplification and DNA sequencing

The PCR product underwent two rounds of purification using MicroconR 100 (Amicon) to remove the unincorporated primers and primer artifacts. An aliquot of 1 to 2 µl was used to perform and half or full nest secondary PCR amplification (12 cycles using the above stated conditions) to attain high product yield without obvious primer artifacts. Two additional rounds of PCR product purification were completed prior to standard sequence analysis using automated fluorescent sequencing (ABI377, PE Applied Biosystems) and Big Dye terminator chemistry with Ampli*Taq* FS DNA polymerase (PE Applied Biosystem). The data were analyzed by SequencherTM software (version 3.1.1, Gene Codes). A mutation was called when the mutant peak volume was at least 20% of the primary wild type peak at the same position in either sequencing direction. Allele dropout occurred when the peak of one allele was less than 10% of the other allele peak at a known heterozygous base position in either sequencing direction.

Year One: Challenges – Solutions

1. Measure individual mutation pattern
   a) identify mutant cells – use of immunohistochemical staining to identify cells with overabundance of p53 protein and with potential for containing a conserved missense mutation in exons 5 to 9 of the p53 gene.
2. maintain high DNA integrity (>20 kb single stranded DNA)
   a) optimize tissue fixation – use of ethanol/EDTA for tissue fixation
   b) optimize antigen retrieval and immunohistochemical staining – use of steam heat for antigen retrieval
   c) optimize tissue section thickness to obtain fully intact nuclei – use of 6 micron tissue sections
   d) retrieve intact, single nuclei – use of micromanipulator-assisted dissection of single cells
3. attain high efficiency single cell PCR – development of Stimulated PCR
4. attain high quality DNA sequencing – use of extensive PCR product purification and a secondary half or full nest PCR amplification
5. achieve a low rate of polymerase error and allele drop out – maintenance of high DNA integrity, use of polymerases with low error rate, acquisition of intact nuclei and use high-efficiency PCR technique.

Solutions were identified for each challenge

Year two: Challenges- solutions

1. Reduce assay cost and increase assay efficiency-incorporated DOVAM-S mutation scanning.
2. Determine how to reduce allele dropout-developed and employed TA.
3. Ensure peer review of assay efficiency and applicability- published novel research developments and findings.

Acceptable solutions were identified for each challenge.

Year Three: Challenges- solutions

The frequency of positively stained with p53 was less than anticipated

Solutions:

I) **This problem was ameliorated in multiple ways. The staining protocol was re-examined and controlled more rigorously. In addition, the focus of analysis was shifted to sections of breast tissue that had a high fraction of epithelial cells relative to fat cells.**

II) **To increase the sensitivity of the immunohistochemical staining we used additional anti-p53 antibody which recognize a.a 213-217 of human P53 (PAB 240). The new antibody was used at a final concentration of 1:150 which is the same concentration used for DO-7. The cocktail of the two antibodies have considerably increased the sensitivity of the staining.**

III) **The DAKO EnVisionTM System (K4016) kit was used to increase sensitivity and**

decrease time of staining. The DAKO EnVisionTM System AP uses an AP labeled polymer which is conjugated to secondary antibodies. This polymer is ready-to-use and eliminates the need for separate applications of link antibody and alkaline phosphatase conjugated reagents. The polymer does not contain avidin or biotin. Consequently, nonspecific staining resulting from endogenous avidin-biotin activity in some tissue is eliminated or significantly reduced.

Relevance

The incidence and mortality from breast cancer varies significantly in ethnically and geographically distinct populations (Boring et al., 1993). Furthermore, when certain low risk populations, particularly Asian women, immigrate to the West there is a gradual increase in the incidence of breast cancer to a level approaching that of the indigenous population (Ziegler et al., 1993). We showed that patterns of genetic damage in the p53 gene, a gene highly associated with cancers of all types, varies depending on ethnicity and/or place of residence (Hartmann et al., 1997). We suspect that this variability is related to differences in exposure to environmental mutagens, presumably through differences in the content of toxic substances in the diet (Biggs et al., 1993; Blaszyk et al., 1996). **The development of this p53 mutation load assay permits determination of a pattern of mutation in an individual.** Mutation patterns will be determined for twenty upper Midwest U.S. women, ten with diets high in animal fat and ten with diets low in animal fat. In each woman, 30 different mutations will be identified in epithelial cells of normal mammary tissue. With this newly developed technique, it is possible to determine the mutation fingerprint in each woman and correlate this with her diet. It is hypothesized that there will be substantial individual variation in mutation patterns within the population consistent with a key prediction of the "lipophilic mutagens hypothesis".

KEY RESEARCH ACCOMPLISHMENTS (months 1 to 12)

1. Development of a protocol with maintenance of high integrity DNA (>20 kb single stranded DNA) following fixation, embedding, immunohistochemical staining, microdissection and digestion.
2. Development of single cell PCR with high efficiency, low polymerase error and low allele drop out (ADO).

3. Determination of % cells stained for p53 overabundance with conserved missense changes.
4. Determination of % unstained cells with conserved missense changes.
5. Measurement of % ADO.
7. Development of a database for tracking samples, and recording and analyzing mutations.

KEY RESEARCH ACCOMPLISHMENTS (months 13 to 24)

1. Incorporation of DOVAM-S mutation scanning to increase assay efficiency and reduce assay cost
2. Development of TA to reduce allele dropout
3. Publication of key research findings

KEY RESEARCH ACCOMPLISHMENTS (months 25 to 36 with no cost extension)

1) **Improving staining time by using polymer technology: improving staining time not only saves time but also help to keep the integrity of the DNA intact by reducing the time in which the tissue is being processed at room temperature. Usually tissue blocks are kept at 4° C until processed and slides at minus at - 80° C until micro-dissected. If out of the freezer the slides are always kept on dry ice.**

2) **Adding a second antibody which recognize an epitope different from the one recognized by DO-7. Using a cocktail of the PAB 240 and the DO-7 with final concentration of 1:150 for each antibody increased the sensitivity of the staining.**

3) **Testing SCIMLA on archival tissues which are collected at different periods. The success of amplification and genetic analysis from these tissues make SCIMLA applicable to wide variety of archival tissues if they were fixed in appropriate fixative and a specific protocol is followed.**

4) **Demonstrating that tissue from normal breast tissue can harbor up to nine missense mutations (table 1). Such number of mutations in one individual from normal tissue is very impressive and has not been demonstrated before. We**

anticipate that the number of mutations will increase as the dissection and analysis is still underway.

5) Adopting dietary questionnaire. After reviewing and studying many questionnaires. The PI decided to use "Food and nutrition questionnaire of the breast cancer comprehensive project".

6) 196 cells positive for p53 were dissected from one case, out which 46 cells have missense mutations. A total of nine unique missense mutation were found in this case so far. Table 1 summarizes the mutations in this patient. Two jackpots mutation were identified one in codon 175 (11 cells) and the other in codon 245. Both mutations are transitions (G>A). 300 hundred neighboring white cells were analyzed as controls. Three have showed the same missense mutation as codon 245 and one showed the same missense mutation as codon 175.

7) Determination of mutation load in normal mammary tissue from three patients (see above).

8) Demonstration that rare cells can be detected in mammary cells by p63's straining; although formal proof is not yet available, the data suggest that p63 will be a useful reporter gene, augmenting in supplementing the data available for p53. Since p63 is rarely mutated in tumors, p63 promises to be a general tool for measuring mutation load in breast cancers.

8. Demonstration that multiplex PAP can be used to amplify multiple exons from single cells using a model system.

OUTCOMES:

Manuscripts

Schlake G, Liu Q, Heinmoller E, Hill KA, Weiss L, Sommer SS. Single-cell immunohistochemical mutation load assay (SCIMLA) using human paraffin-embedded tissues. Environ Mol Mutagen. 2003;42(3):206-15.

Heinmoeller, et al., 2001. **Toward efficient analysis of mutations in single cells from ethanol-fixed, paraffin-embedded and immunohistochemically-stained tissues.** Laboratory Investigations 82(4): 443-53,2002.

Hill and Sommer, 2001. **P53 as a mutagen test in breast cancer.** Environmental and Molecular Mutagenesis 39: 216-227,2002.

Abstracts

Sommer et al., 2004. **SCIMLA, A general tool for measuring somatic mutation load in solid tissues: Dramatic inter-individual variation mutagen hypothesis of breast cancer.** Submitted for American Association for Cancer Research.

Sommer et al., 2001. **Measuring mutation load in human cells from paraffin-embedded human tissues.** Annual Meeting of the Environmental Mutagen Society (poster presentation).

Hill K.A. 2001. **P53, ATM and breast cancer: A keynote address.** Breast cancer and environmental mutagens: Bridging molecular research to medicine and public health. A special interest group meeting of the Environmental Mutagen Society.

Schlake et al., 2001. **Measurement of mutation load using microdissected single cells from ethanol-fixed, paraffin-embedded and immunohistochemically-stained human tissues.** Annual meeting of the American Society of Human Genetics (poster presentation).

Meetings

2004 Annual Meeting of the American Society of Human Genetics

2004 Annual Meeting of the Environmental Mutagens Society

2004 The Jackson Laboratory "Mutagenesis Meeting"

2003 Annual Meeting of the Environmental Mutagens Society

2003 Annual Meeting of the American Society of Human Genetics

2002 Annual meeting of the environmental Mutagen Society

2001 Annual meeting of the Environmental Mutagen Society.

2001 Breast cancer and environmental mutagens: Bridging molecular research to medicine and public health. A special interest group meeting of the Environmental Mutagen Society.

2001 Annual meeting of the American Society of Human Genetics.

Presentations

Sommer, Steve S. 2002. Human Mutation Load Assay(HMLA) using the Tp53 gene of single cells from paraffin-embedded human tissues. The annual meeting of the environmental Mutagen Society.

Hill, 2001. Measurement of mutation load using the p53 gene in human cells from paraffin-embedded tissues. The annual meeting of the Environmental Mutagen Society; Special Interest Group Session on Novel Technologies.

Hill K.A. 2001. **P53, ATM and breast cancer: A keynote address.** Breast cancer and environmental mutagens: Bridging molecular research to medicine and public health, a special interest group meeting of the Environmental Mutagen Society.

CONCLUSIONS

The newly developed p53 mutation load assay permits detection of mutations in 2 kb of the p53 gene amplified from single cells immunohistochemically stained for overabundance of p53 protein and manually microdissected from paraffin-embedded, normal human tissues. The assay identified conserved missense changes in 2 kb of the p53 gene in single cells with over expression of the p53 protein. Rarely do cells with normal p53 over expression have conserved missense changes in the p53 gene. The rate of allele drop out in the performance of this assay is

low. The probability of identifying rare heterozygous mutations is high. The method is applicable to a variety of human tissues that are ethanol fixed, paraffin embedded and immunohistochemically stained. Careful DNA preservation, sensitive immunohistochemical staining, precise manual microdissection of single cells and high efficiency amplification from single cells allows mutation load to be measured in single cells of normal human tissues. We demonstrate detection of rare heterozygous mutations in at least 2 kb of the p53 gene amplified from single cells. **It has now become possible to measure mutation load in an individual by identifying p53 mutations in the normal mammary cells in an individual. This new tool greatly increases the power to test the central hypothesis. The original tasks proposed have been revised to permit measurement of mutation pattern in an individual.** We propose to test twenty upper Midwest U.S. women with breast cancer, ten with diets high in animal fat and ten with diets low in animal fat. In each woman, 30 different mutations will be defined in mammary cells. The normal breast tissue is obtained away from the tumor margin in women with breast cancer. With this newly developed technique, it is possible to determine the mutation fingerprint in each woman and correlate this with her diet. It is hypothesized that there will be substantial individual variation in mutation patterns within the population consistent with a key prediction of the "lipophilic mutagens hypothesis".

REFERENCES

Biggs, P.J., Warren, W., Venitt, S. and Stratton, M.R. Does a genotoxic carcinogen contribute to human breast cancer? The value of mutational spectra in unravelling the aetiology of cancer. *Mutagenesis* 8:275-283, 1993.

Blaszyk, H., Hartmann, A., Liao, D., Kovach, J.S. and Sommer, S.S. Evidence for diverse mutagens in breast cancer. *Lancet* 348:683-684, 1996.

Boring, C.C., Squires, T.S. and Tong, T. Cancer statistics, 1993. *Ca: a Cancer Journal for Clinicians* 43:7-26, 1993.

Boring, C.C., Squires, T.S., Tong, T. and Montgomery, S. Cancer statistics, 1994. *Ca: a Cancer Journal for Clinicians* 44:7-26, 1994.

Hartmann, A., Blaszyk, H., Kovach, J.S. and Sommer, S.S. The molecular epidemiology of *p53* gene mutations in human breast cancer. *Trends Genet.* 13:27-33, 1997.

Ziegler, R.G., Hoover, R.N., Pike, M.C., et al. Migration patterns and breast cancer risk in Asian-American women. *J.Natl.Cancer Inst.* 85:1819-1827, 1993.

APPENDICES

Manuscripts

1. Schlake G, Liu Q, Heinmoller E, Hill KA, Weiss L, Sommer SS. Single-cell immunohistochemical mutation load assay (SCIMLA) using human paraffin-embedded tissues. Environ Mol Mutagen. 2003;42(3):206-15.

2. Heinmoeller, et al., 2001. Toward efficient analysis of mutations in single cells from ethanol-fixed, paraffin-embedded and immunohistochemically-stained tissues. Laboratory Investigations 82(4): 443-53,2002.

3. Hill and Sommer, 2001. P53 as a mutagen test in breast cancer. Environmental and Molecular Mutagenesis 39: 216-227,2002.

Abstracts

1. Sommer et al., 2004. **SCIMLA, A general tool for measuring somatic mutation load in solid tissues: Dramatic inter-individual variation mutagen hypothesis of breast cancer.** Submitted for American Association for Cancer Research.

Environmental and Molecular Mutagenesis 42:206–215 (2003)

Single-Cell Immunohistochemical Mutation Load Assay (SCIMLA) Using Human Paraffin-Embedded Tissues

Gudrun Schlake,[1] Qiang Liu,[2] Ernst Heinmöller,[1] Kathleen A. Hill,[1] Lawrence Weiss,[3] and Steve S. Sommer[1,2]*

[1]*Department of Molecular Genetics, City of Hope National Medical Center and Beckman Research Institute, Duarte, California*

[2]*Department of Molecular Diagnosis, City of Hope National Medical Center and Beckman Research Institute, Duarte, California*

[3]*Department of Pathology, City of Hope National Medical Center and Beckman Research Institute, Duarte, California*

It would be advantageous to measure mutation load in situ in order to determine the relationship between a high mutation load and increased risk for cancer or other diseases and to evaluate sources of possible mutagen exposure. Previously, in situ mutation detection assays have been plagued with multiple rounds of amplification and high rates of false-positives and false-negatives. The single cell immunohistochemical mutation load assay (SCIMLA) was developed to measure somatic mutation frequency, pattern, and spectrum in normal tissues with a single round of amplification. The *P53* gene was utilized as a mutation reporter because of the unusual property that missense mutations often cause P53 protein to accumulate in the cell, allowing the mutant proteins to be detected by immunohistochemical staining. Alternative reporter genes with stabilized mutant proteins may be envisioned. Single cells that stain positively for P53 protein overabundance (red cells) were microdissected from ethanol-fixed and paraffin-embedded tissues. A novel stimulated-PCR (S-PCR) protocol permitted successful amplification of a 1.8-kb segment of the *P53* gene (i.e., exons 5–9) in 87% of single mammary cells. Subsequent sequence analysis demonstrated that 35% of the amplified red-stained epithelial cells from normal breast tissue have missense mutations at evolutionarily conserved amino acids. Jackpot mutations, presumably due to clonal expansion, were common. False-positive missense mutations at conserved residues were observed in 3% of the clear cells (i.e., without red stain), presumably due to DNA polymerase error in early PCR cycles. The allele dropout rate was measured at 40% of the amplified cells. SCIMLA is applicable to a variety of tissues, utilizes a single amplification of an endogenous gene, displays mutant cells in situ, and may be adapted to other species. Environ. Mol. Mutagen. 42:206–215, 2003. © 2003 Wiley-Liss, Inc.

Key words: *P53* gene; single cell analysis; single cell microdissection; single cell amplification; mutation load; allele dropout; immunohistochemical staining

INTRODUCTION

Genetic damage accumulates over time as a result of endogenous factors or environmental insults. Detection of somatic mutations within single cells would permit measurement of an individual's mutation load (i.e., the mutation frequency, pattern, and spectrum). These measurements in normal tissues can identify individuals with a high mutation load at increased risk for disease, including cancer, by indicating an endogenous predisposition or a mutagen exposure risk.

Mutation load has been studied in rodents using transgenic mouse mutation detection systems and the best studied contain either the *lacI*, *lacZ*, or *supF* genes integrated within a mouse chromosome [Hill et al., 1999; Heddle et al., 2000; Nohmi et al., 2000; Lewis et al., 2001]. Data now exist on the relationship of mutation load to the life cycle of the mouse [Stuart et al., 2000; Ono et al., 2000], endogenous defects [Andrew et al., 1996, 1997, 2000; Buettner et al., 1997], infection [Motorna et al., 2001], and mutagen exposure [Skopek et al., 1995, 1996; Walker et al., 1996]. These data are averages for all the cell types in a tissue, unless specific cell types can be purified.

Mutation load can be measured in T-cells by utilizing the

© 2003 Wiley-Liss, Inc.

Grant sponsor: DAMD; Grant number: 17-00-1-0204.

*Correspondence to: Steve S. Sommer, M.D., Ph.D., Director, Departments of Molecular Genetics and Molecular Diagnosis, City of Hope National Medical Center, 1500 East Duarte Road, Duarte, CA 91010-3000. E-mail: sommerlab@coh.org

Received 23 May 2003; provisionally accepted 21 June 2003; and in final form 10 July 2003

DOI 10.1002/em.10193

HPRT gene [Recio et al., 1990; Cariello and Skopek, 1993; Cole and Skopek, 1994; Albertini, 2001]. Recently, the approach was extended to kidney [Colgin et al., 2002]. Mutation load has been measured at four nucleotides of the *P53* gene [Aguilar et al., 1993; Hussain et al., 2000, 2001] by using a mutation enrichment method with multiple rounds of restriction endonuclease digestion and PCR amplification. A high mutation load was observed in normal lung tissue in individuals with lung cancer and *P53* hotspot codons showed the signature mutations for known mutagens in cigarette smoke [Hussain et al., 2001]. Increased *P53* mutation load in normal colon tissue was observed in individuals with ulcerative colitis, a cancer-prone chronic inflammatory disease [Hussain et al., 2000]. This assay is limited to a few nucleotide substitutions and does not permit in situ identification of mutant cells, determination of mutant frequency, or involve the molecular analysis of single cells.

A whole-mount preparation method examined mutation load in situ in clusters of 60–3,000 cells of the human epidermis using immunohistochemical analysis of stabilized P53 protein and detected frequent clonal expansion of *P53*-mutated keratinocytes [Jonason et al., 1996]. Molecular analysis of the mutant clones revealed the mutation signature of UV in sun-exposed skin in which P53-stained cells cluster and are much more abundant than in other tissues. It is desirable to identify a method for in situ detection of single cells with endogenous gene mutations for measurement of in vivo mutation load.

The *P53* gene has several properties that may make it an advantageous reporter for single nucleotide substitutions or short in-frame deletions: 1) The region of likely functional significance is small enough to allow rapid scanning, i.e., *P53* mutations are clustered within exons 5–9 (in breast cancer, 93% of mutations occur within exons 5–9 [Hartmann et al., 1995]); 2) the P53 protein has the unusual property that missense mutations often stabilize the protein, permitting immunohistochemical staining to identify mutant *P53* cells [Bartek et al., 1990]; 3) substantial information is available on the level of *p53* sequence conservation [Soussi et al., 1990; Walker et al., 1999]; and 4) the frequency and pattern of mutation are generally unchanged in normal tissues and elevated little, if at all, in tumor tissues of *p53*-nullizygous mice [Sands et al., 1995; Buettner et al., 1996, 1997], suggesting that the background rate of mutation in p53-deficient cells is similar to that of the wild-type cells.

Previously, we reported progress toward analysis of mutations in single cells from ethanol-fixed, paraffin-embedded, and immunohistochemically stained tissues [Heinmoller et al., 2002a]. Single cells from ethanol-fixed and paraffin-embedded tissues that stain positive for P53 overabundance were microdissected from normal human tissues and exons 5–9 of the *P53* gene were amplified individually to measure mutation load based on detection of primarily missense mutations in exons 5–9 of the *P53* gene that lead to overabundance of P53 protein detectable by immunohistochemical staining. The method used a two-step amplification protocol composed of whole genome amplification using a modified improved primer extension preamplification–PCR, followed by specific multiplex nested PCR. The protocol was cumbersome and suffered from a high rate of allele dropout (ADO, i.e., detection of only one of two possible alleles), such that at least one of exons 5–9 were not amplified in more than 85% all patients cells.

Herein, the single cell immunohistochemical mutation load assay (SCIMLA) was developed to measure in situ the frequency, pattern, and spectrum of rare somatic *P53* mutations using a novel 1.8-kb single-step amplification protocol specifically designed for amplification from single cells with minimal PCR-associated ADO. A single robust amplification of exons 5–9 produces a tolerable rate of ADO and low rate of false-positive mutations. An important feature of the assay is that individual mutant cells can be placed in the architecture of intact tissues.

MATERIALS AND METHODS

Sample Acquisition

Normal tissues from breast, colon, and lung were obtained from cancer patients undergoing surgery at the City of Hope National Medical Center. When the patient had a tumor, the normal tissue, as judged by a pathologist (GS), was dissected away from the tumor margin. The study has IRB approval.

Tissue Processing and Immunohistochemical Staining

The tissue samples were cut into 3–4 mm slices, transferred immediately into jars containing an ethanol-based fixative (85% ethanol with 0.2 mM EDTA buffer, pH 8.0), and incubated for 3 days. The specimens were embedded in paraffin. Serial sections of 6 μm thickness were cut. Deparaffinization included one xylene step at room temperature for 30 min with shaking every 5 min, followed by serial rehydration using 100%–70% ethanol. Shaking is necessary for complete deparaffinization and a high rate of amplification. Antigenic sites of the P53 protein were unmasked using steam heat for 5 min in 1 mM EDTA buffer (pH 8.0) at 96–100°C. The mouse monoclonal p53 antibody DO7 (Novocastra, Newcastle Upon Tyne, UK) was used for immunohistochemical staining at a titer of 1:200. The tissue sections were stained immunohistochemically for overabundance of P53 protein and counterstained with hematoxylin. Cells with positive nuclear staining for an overabundance of P53 protein were termed red cells and had bright red nuclei. In contrast, clear cells showed no P53 antibody staining and had hematoxylin (blue)-stained nuclei.

Single Cell Microdissection

Single cells were microdissected using an inverted microscope (Nikon TMS, Tokyo, Japan) and a mechanical micromanipulation system (Model MP-285 from Sutter Instruments, Novato, CA; protocol reviewed in Heinmoller et al. [2002b]). Under direct visualization of the immunohistochemical staining, a joystick was used to manipulate a tungsten needle to separate single cells from the surrounding tissue. The freed single cell with a well-preserved nucleus was manually transferred with a sterile 27 G1/2″ needle into a 0.2 ml PCR tube containing 5 μl digestion buffer (#3 High

TABLE I. List of Primers for the *P53* Gene

	Primer					PCR segment[c]		
#	Name[a]	Sequence (5′-3′)	Size (base)	T_m (°C)[b]	GC%	Size (bp)	Tm (°C)[2]	GC %
1	P53(13005)30D	TGTTCACTTGTGCCCTGACTTTCAACTCTG	30	65	46.7	1858	81.7	54.5
2	P53(14862)30U	CCTGATGGCAAATGCCCCAATTGCAGGTAA	30	71.5	50.0			
3	P53(13007)30D	TTCACTTGTGCCCTGACTTTCAACTCTGTC	30	64.3	47.7	1848	81.7	54.5
4	P53(14853)30U	AAATGCCCCAATTGCAGGTAAAACAGTCAA	30	66.2	40.0			
5	P53(13015)24D	TGCCCTGACTTTCAACTCTGTCTC	24	56.4	50		D sequencing	
6	P53(13280)20D	AGGGTCCCCAGGCCTCTGAT	20	58.3	65		D sequencing	
7	P53(13489)22U	GGCCACTGACAACCACCCTTAA	22	57.5	54.5		U sequencing	
8	P53(13949)20D	AGGTCTCCCCAAGGCGCACT	20	59.5	65		D sequencing	
9	P53(14153)20U	GGGGCACAGCAGGCCAGTGT	20	61.7	70		U sequencing	
10	P53(14505)19D	GGAGAGACCGGCGCACAGA	19	58.4	68.4		D sequencing	
11	P53(14815)24U	CGGCATTTTGAGTGTTAGACTGGA	24	57.1	45.8		U sequencing	

[a]Informative names give the precise sizes and locations of the PCR fragment. As an example for P53(13005)30D, P53 = the *P53* gene, (13005)30D = 5′ end of the primer begins at 13,005 base position, and the length is 30 bases "downstream" (D) (i.e., in the direction of transcription). The sequence of the *P53* gene was from a revised version of X54156 in GenBank.

[b]T_m of the primer was estimated by the nearest-neighbor method at 50 mM KCl and 250 pM DNA and T_m of the PCR segment was estimated by the formula of Wetmur: $T_m^{product} = 81.5 + 16.6 \log[K^+ = 0.05\ M] + 0.41(\%G + \%C) - 675/length$.

[c]Primer pair of 1 and 2 was used for S-PCR; primer pair of 3 and 4 was for nested PCR; Primers 5 to 11 were for sequencing.

Fidelity buffer without Mg^{++}; Roche, Indianapolis, IN), 2 mg/ml Proteinase K (Qiagen, Santa Clarita, CA), 3% Tween-20 detergent, and 0.2 mM EDTA, pH 8.0. The single cell was digested at 50°C for 16 hr followed by inactivation of Proteinase K by heating at 90°C for 10 min.

Length of the Single-Stranded DNA Template

To estimate the average length of the single-stranded genomic DNA following tissue processing, 3,000 cells were scraped from the tissue section and digested. The genomic DNA was denatured by heating and electrophoresed through a 0.8% agarose gel to determine the average length, representing most likely the length of the single-stranded DNA template from a single cell.

Stimulated-PCR (S-PCR) From Single Cells

The 1.8-kb segment including exons 5–9 of the *P53* gene was amplified from single cells using S-PCR after systematically evaluating multiple parameters. The PCR components included exogenous carrier mouse DNA, bovine serum albumin, primer, antibody-inactivated Platinum High Fidelity *Taq* DNA polymerase (*Taq/GB-D* DNA polymerases), Mg^{++}, dNTP, deaza-dGTP, and DMSO; and the thermocycling conditions, including annealing and extension temperatures and periods, were explored extensively. For example, high concentrations of the enzymes (2.5 U of *Taq/GB-D* polymerases/25 μl reaction volume), 4-fold the concentration recommended by the manufacturer (Invitrogen, Carlsbad, CA), achieved the highest amplification yield. Mixing of Platinum *Taq* with Platinum *Taq/GB-D* High Fidelity DNA Polymerases in 1:1 unit ratio, which increased the unit ratio of *Taq* to *GB-D* by 2-fold, amplified better than the High Fidelity enzymes alone, indicating that not only the total units but also the ratio of the two enzymes are important. Hot-start using *Taq* antibody to inactivate *Taq* DNA polymerase at room temperature generated fewer primer artifacts using 40–45 cycles. Primers with AA dinucleotides at the 3′ end also formed fewer dimers [Gelfand et al., 1999].

Primer design was critical. Primers were designed with strict criteria to ensure specificity and efficiency (Table I). The successful pair of oligonucleotides were designed and analyzed with Oligo 5 software (National Biosciences, Plymouth, MN). The T_m of the primer was estimated by the nearest-neighbor method at 50 mM KCl and 250 pM DNA and the T_m of the PCR segment was estimated by the formula of Wetmur [1991]: $T_m^{product} = 81.5 + 16.6 \log[K^+ = 0.05\ M] + 0.41(\%G + \%C) - 675/length$. The criteria for specificity included: 1) a maximum of three complementary nucleotides at the 3′ ends of each oligonucleotide; 2) no false priming sites on the mouse *p53* gene (due to the presence of mouse genomic DNA during S-PCR); and 3) no false priming sites on the human *P53* gene, i.e., no more than seven bases at the 3′ end of the primers were identical to the 1.8-kb amplified segment. The *P53* gene from single cells was amplified with the "PCR Heaven" protocol used for a robust, simple, and quantitative PCR protocol [Liu et al., 2003]. "PCR Heaven" is a term that refers to high levels of key components of PCR (concentrations of primer, DNA polymerases, dNTPs, additives, and magnesium optimized in the presence of mouse genomic DNA carrier), such that amplification is virtually 100% efficient per cycle under normal conditions, although the efficiency for one or two copies of genome template is unclear.

The PCR mixtures contained in a total volume of 25 μl: human genomic DNA from a single microdissected cell, #3 Expand High Fidelity buffer (Roche), 3.5 mM $MgCl_2$; 500 μM of each dNTP, 2% DMSO; 0.4 μM of each primer, and 5 μg of BSA. A mixture containing a very high polymerase concentration (1.25 U of Platinum *Taq/GB-D* High Fidelity DNA Polymerases / 1.25 U of Platinum *Taq* DNA polymerase (Invitrogen)) and 25 ng of mouse genomic DNA was critical to success. The cycling conditions included denaturation at 92°C for 12 sec, annealing at 60°C for 20 sec, and elongation at 68°C for 2 min for 40 or 45 cycles with a Perkin Elmer GeneAmp PCR system 9700 (Boston, MA). An additional 20 sec of denaturation preceded the first cycle. As a positive control for S-PCR, a sample of 50–100 cells was scraped from the slide. Aliquots of 2 μl of the PCR product were electrophoresed through a standard 1% agarose gel and the gel was stained with ethidium bromide for UV photography by a CCD camera (Bio-Rad Gel Doc 1000; Hercules, CA). The PCR product was purified using Microcon® 100 columns (Amicon, Bedford, MA) to remove the dNTPs, unincorporated primers, and primer artifacts.

Sequence Analysis and Mutation Identification

A second nested PCR was performed for 12–15 cycles under standard conditions of Platinum *Taq/GB-D* High Fidelity DNA Polymerases to obtain higher product yield. The PCR product purification was repeated as given above. DNA sequencing was performed using an ABI 377 fluores-

cent DNA sequencer and BigDye terminator chemistry with Ampli*Taq* FS DNA polymerase (Applied Biosystems, Foster City, CA). The sequencing primers are listed in Table I. The data were analyzed using Sequencher™ sequence analysis software (v. 3.1.1; Gene Codes, Ann Arbor, MI). Chromatogram primary and secondary peak areas were computed using Phred [Ewing and Green, 1998; Ewing et al., 1998]. A mutation was identified as a second peak that had at least 20% of the wild-type peak area at the same base position in one of the sequencing directions. ADO was identified when the peak area of one allele was less than 20% of the other allele at a known heterozygous base position.

Contamination

Contamination can be a significant problem for amplification from single cells, because of the precision required for microdissection of single cells from tissue sections, limited target template, various sources of other DNAs, and extensive PCR amplifications. The following steps were taken to minimize potential contamination: 1) negative controls without target DNA (prepared by transferring scrapings from regions of the tissue section not containing nuclei) were always used, and 2) the microdissection and preparation of amplification from single cells were always performed in a "clean room" (i.e., isolated from post-PCR sample processing and analysis).

RESULTS

Preliminary Estimates of Mutant Frequency Based on Immunohistochemical Staining

The average number of "positive" red cells per tissue section was examined for normal tissue found in four breast cancer patients. The average staining frequency was estimated at 2.5×10^{-4}.

S-PCR From Single Cells

A detailed analysis of amplification parameters led to robust amplification of a 1.8-kb segment with exons 5–9 of the *P53* gene from single cells. The S-PCR protocol was successful in 87% of amplifications of the *P53* gene from single mammary cells. A typical S-PCR amplification is shown in Figure 1, Step IV. Among the more important parameters for success with single cells are appropriate design of primers, addition of mouse DNA, high concentration of two DNA polymerases, and the hot-start method (see Materials and Methods).

Mutations Identified in Stained Single Cells

Forty-three red cells were microdissected from four patients. The 1.8-kb region of the *P53* gene was amplified and 1.6 kb was sequenced (including exons 5–9). A total of 21 sequence changes were observed (49%), including 16 missense changes (Table II). Of the missense changes, 15 occurred at evolutionarily conserved amino acids (35%) and were scored as bona fide somatic mutations associated with the red staining. Two jackpot mutations were observed (Table II; 13419A→G, n = 9 cells from individual 2 and 13155C→T, n = 4 cells from individual 3). Each mutation was found in only one person and each mutation cosegregated with the same nearby polymorphic allele, observations consistent with clonal expansion. Sequence analysis of the tumors indicates that these jackpot mutations were *not* in the clones that gave rise to the breast cancer in these women. A significantly greater frequency of conserved missense mutations was observed in red compared to clear cells ($P < 0.0001$; Fisher's exact test).

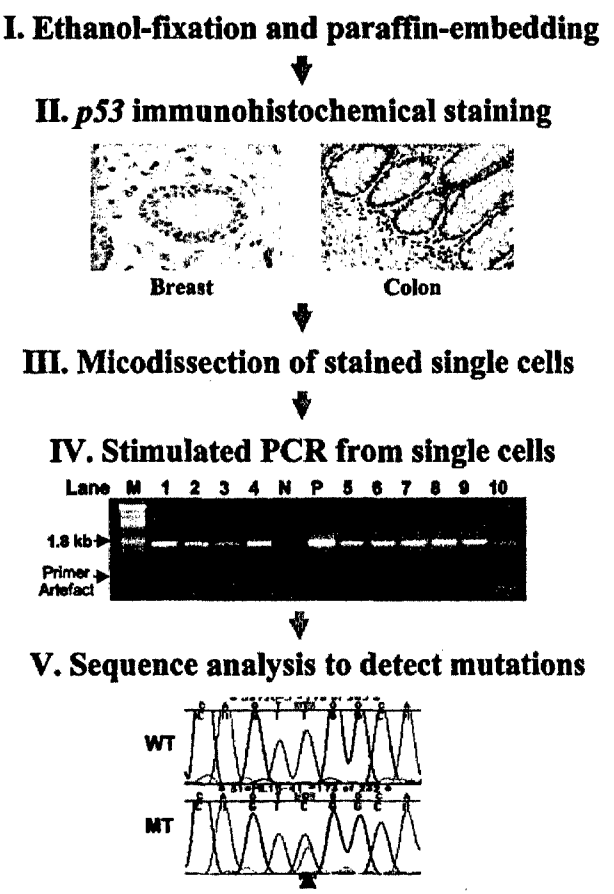


Fig. 1. A flow chart of the single cell immunohistochemical mutation load assay (SCIMLA) protocol. I. Modified tissue processing maximizes recovery of intact nuclei and genomic DNA. II. Immunohistochemical staining of normal breast and colon tissue identifies mutant cells with a frequency of 2.5×10^{-4}. Bright red nuclei indicate positive staining for overabundance of P53 protein. Clear cells do not show immunohistochemical staining for P53 overabundance. III. Microdissection of stained single cells selects nuclei with a high fraction of the total nuclear volume. IV. Stimulated-PCR (S-PCR) is successful in 87% of single cells. A 1.8-kb region of the *P53* gene is amplified from genomic DNA of microdissected single cells using 40 cycles. Lanes 1–10 are single-cell samples. Lane N = no DNA control. Lane P = positive control with 20 cells, M = DNA marker. The S-PCR product (1.8 kb) and primer artifacts (200 bp) are indicated. V. Sequence analysis to detect mutations shows a typical pattern of the wild-type (WT) and a heterozygous mutation (MT) from single cells.

TABLE II. Sequence Changes in the *P53* Gene in Single Cells

Individual ID[a]	Cell ID	Sequence change[b]	Amino acid change	Species conservation[c]					
				Monkey	Mouse	Rat	Chicken	Xenopus	Trout
	Red Cells								
1	1	G14617A	NC[d]						
1	2	A14582G	Lys ⇒ Arg	Y	Y	Y	Y	Y	Y
2	3	A13419G	Tyr ⇒ Cys	Y	Y	Y	Y	Y	Y
2	4	G13238A	Gly ⇒ Ser	Y	Y	Y	Y	N	Y
2	5	A13419G	Tyr ⇒ Cys	Y	Y	Y	Y	Y	Y
2	6	A13245G	NC[d]						
2	6	T14675C	NC[d]						
2	7	A13419G	Tyr ⇒ Cys	Y	Y	Y	Y	Y	Y
2	8	A13419G	Tyr ⇒ Cys	Y	Y	Y	Y	Y	Y
2	9	A13419G	Tyr ⇒ Cys	Y	Y	Y	Y	Y	Y
2	10	A13419G	Tyr ⇒ Cys	Y	Y	Y	Y	Y	Y
2	11	A13419G	Tyr ⇒ Cys	Y	Y	Y	Y	Y	Y
2	12	C14793T	NC[d]						
2	13	A13419G	Tyr ⇒ Cys	Y	Y	Y	Y	Y	Y
2	13	A14067G	Glu ⇒ Gly	Y	Y	Y	Y	Y	Y
2	14	A13419G	Tyr ⇒ Cys	Y	Y	Y	Y	Y	Y
3	15	C13155T	Ala ⇒ Val	Y	Y	Y	Y	Y	Y
3	16	C13155T	Ala ⇒ Val	Y	Y	Y	Y	Y	Y
3	17	C13155T	Ala ⇒ Val	Y	Y	Y	Y	Y	Y
3	18	C13155T	Ala ⇒ Val	Y	Y	Y	Y	Y	Y
4	19	T13377G	Leu⇒Trp	Y	N	Y	Y	N	N
	Clear cells								
1	1	G13637A	NC[d]						
2	2	C13984T	NC[d]						
3	3	C13253T	NC[d]						
3	4	T13413C	Val ⇒ Ala	Y	Y	Y	Y	N	N
3	4	C14283T	NC[d]						

[a]Data are from normal breast tissue obtained beyond the tumor margin in four women who underwent surgery for removal of a primary breast tumor.
[b]The genomic sequence of the *P53* gene was from a revised version of X54156 in GenBank. The sequence changes were not present in the germline of these individuals as determined by sequencing of DNA isolated from 1,000 cells.
[c]Species conservation was determined according to Soussi et al. [1990]; a conserved missense change is defined as identity in monkey, mouse, rat, and chicken.
[d]NC = no known functional or structural change.

False-Positive Sequence Changes Likely Due to Polymerase Error

Of the 34 clear cells expected to *not* contain mutations, five sequence changes, including one conserved missense change (3%), were identified in four different cells (Table II). Two of the five changes, including the one conserved missense mutation, occurred in one cell.

Allele Dropout

ADO in the *P53* gene was measured using 48 cells of normal human breast tissue with a known A/G heterozygous polymorphism. The ratio of the area of the A peak to the G peak was measured for each cell. ADO was found in 19 of the 48 cells (40%). ADO was scored if the area of a minor peak was less than 20% of the other peak in both sequencing directions (Fig. 2).

Extension of SCIMLA to Other Tissues

In principle, SCIMLA could be used to measure mutation load in many tissues or cell types. Normal colon and lung cells were stained and analyzed as described for breast cells. Nineteen red cells in colon were amplified and nine contained missense changes at conservative amino acid codons, two red cells in lung were examined and two conservative missense mutations were found (not shown). When 48 clear cells were analyzed in multiple tissues, predominately from colon, two conservative missense mutations were found (4%).

DISCUSSION

We found that missense mutations in the *P53* gene can be detected by in situ immunohistochemical staining of normal cells and analyzed by single-cell amplification followed by sequencing. SCIMLA has been successfully applied to single cells dissected from normal breast, colon, and lung

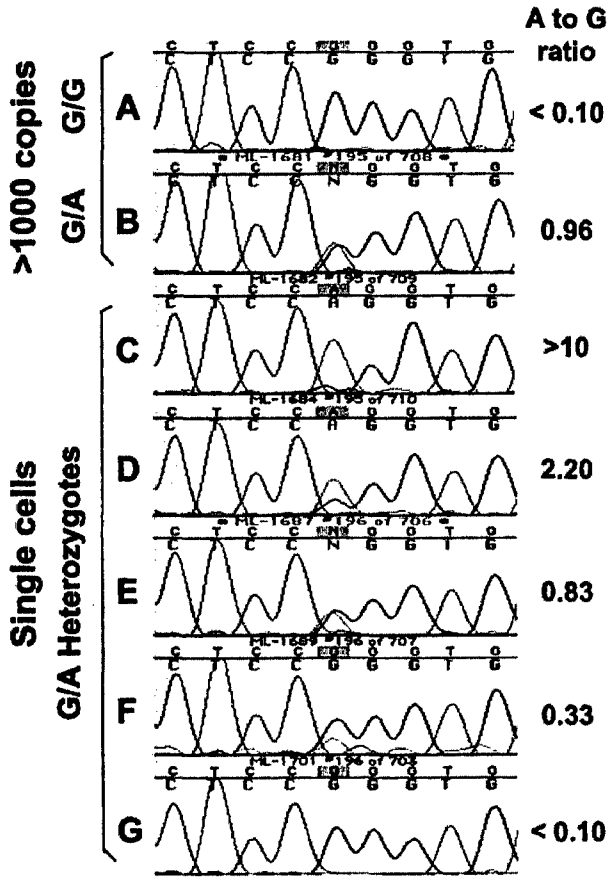


Fig. 2. Peak patterns in a known heterozygous G→A polymorphism. Single cells containing a known heterozygous G→A polymorphism in exon 5 at position 13494 were analyzed for allele dropout (ADO). The amplified product was sequenced using the *P53*(13016)24D primer (Table I). **A:** Sequence from a product that was amplified from more than 1,000 copies of the wild-type DNA. **B:** The positive control, i.e., the heterozygous G→A polymorphism amplified from more than 1,000 genome copies obtained from the same tissue section. **C–G:** The different patterns of the heterozygous G→A polymorphism amplified from single cells. The ratio of G to A peak area is indicated. The peak areas of the G and A nucleotides vary among the single cells.

tissue. About one-third of the stained mammary cells and about one-half of the stained colon and lung cells had *P53* missense changes at conserved amino acids. The data are virtually unchanged if all missense mutations are accepted. Jackpots were observed. The rate of allele dropout was 40%. The rate of false-positives due to polymerase error was detectable but low (3–4%). An important caveat of the assay is the reliance on immunohistochemical detection of elevated P53 protein with detection of heterozygous missense changes in exons 5–9 of the *P53* gene. It is now possible to examine mutation load in individuals as a function of gender, ethnicity, age, diet, mutagen exposure, and genetic predisposition to human diseases, including cancer.

Generally, cells with normal morphology are assayed. Substantial chromosomal abnormalities typically occur later in oncogenesis and it is possible that cells with *P53* mutations would often be polyploid for *P53*, increasing the target size for missense mutations per cell, and elevating the rate of mutation per *P53* gene copy. However, this is balanced by the reduced likelihood of detecting the mutation when the single cell is amplified. If the cell becomes polyploid prior to mutation of a copy of the *P53* gene, the likelihood of finding the mutation is reduced due to a wild-type-to-mutant ratio of substantially greater than 1:1. It seems unlikely that polyploid cells were found and dissected since only cells with normal morphology were analyzed.

SCIMLA Protocol Challenges

The challenge in developing SCIMLA lay in the integration of a multistep protocol in which no individual step was absolutely novel. Ethanol was substituted for standard formaldehyde fixation because it is a coagulant fixative that preserves high-molecular-weight, native DNA with high-quality immunohistochemical staining. Ethanol concentrations were evaluated and EDTA was added to preserve DNA integrity in paraffin blocks. The thickness of the tissue section was evaluated to balance the quality of staining with nuclear integrity; a thickness of 6 μm was chosen. Standard 4-μm sections yield poorer results due to lower amplification rates and higher ADO rates on the successfully amplified products. The thickness of the section needs to be significantly greater than the diameter of the nucleus of the cell type to be analyzed (see discussion on ADO). Conditions for immunohistochemical staining, including the method of antigen retrieval, were evaluated to minimize the number of false-positive, red cells and to preserve high-molecular-weight DNA. The earlier protocols for antigen presentation with microwaves substantially reduce the length of single-strand DNA. Theoretical calculations indicate that the ADO rate for a 1.8-kb segment is at least 45% if the length of the DNA template is 5 kb, while it decreases to 8% if the DNA template length is 20 kb, as herein (data not shown). Immunohistochemical double-staining with PCNA, the product of a gene upregulated with overexpression of normal P53 protein, was utilized to reduce the levels of false-positive staining [Heinmoller et al., 2002a], but this added step was found not to increase the observed mutation frequency.

The microdissection of single cells was performed using a micromanipulator because evaluation of laser-capture microscopes available in 1999 revealed that unacceptable DNA damage occurred to individual cells. If newer models preserve DNA length sufficiently, laser-capture microscopes may substantially enhance the speed of the microdissection step, but manual microdissection does not consume a large fraction of the *total* effort of SCIMLA. In addition, the micromanipulator costs 10% that of the laser-capture microscope. In an attempt to avoid ADO and to increase the success rate of amplification, a preamplification

step (I-PEP) was evaluated [Dietmaier et al., 1999; Heinmoller et al., 2002a], but S-PCR without preamplification was found to be superior.

A threshold effect of template concentration on PCR efficiency was observed. When the template amount was decreased to ≤20 cells within a 25-μl reaction, no amplification was observed. We hypothesized that an important component of the threshold effect was exhaustion of primer supplies due to generation of primer artifacts in the absence of sufficient template. When primers were designed to avoid this and mouse genomic DNA was added, amplification was dramatically improved. Amplification was further optimized by systematically varying other parameters of amplification.

SCIMLA Protocol Improvements

In a preliminary report, we described a prototypical assay in which *individual exons* of the *P53* gene were preamplified with I-PEP and P53, PCNA double-stained cells amplified from ethanol-fixed, paraffin-embedded tissues [Heinmoller et al., 2002a]. One significant problem with the methodology was the high rate of ADO such that one or more exons were missing from >85% of cells. Herein, multiple aspects of the methodology have been improved. Most importantly, an 87% success rate was achieved for amplification of a 1.8-kb segment of the *P53* gene (including exons 5–9) from single cells using the S-PCR protocol. S-PCR and other improvements in tissue processing and immunohistochemical staining provide a reliable and efficient method that dramatically improved the PCR success rate and lowered the rate of both ADO and false-positive mutations.

S-PCR Improves Amplification From Single Cells

It is hypothesized that the addition of mouse genomic DNA contributed to successful amplification by inhibiting template adsorption to tube surfaces, protecting against minimal DNAase activity, activating DNA polymerase by binding directly to the polymerase, stimulating directly extension, and/or facilitating a coiled to stretched conformational transition due to increased DNA concentration [Mitnik et al., 1995]. The yield of S-PCR was sufficient for sequence analysis in most cases, but generated a higher sequencing background, presumably because of a spurious product of about 200 bases that was not removed by purification. A nested PCR eliminated the spurious product. Previously, a 140-bp PCR segment was amplified from a single genome copy using 60 cycles [Vogelstein and Kinzler, 1999]. To our knowledge, this is the first report of a single round of PCR for 40 cycles on DNA from a single microdissected cell that can achieve an amplification product visualized by standard ethidium-bromide staining.

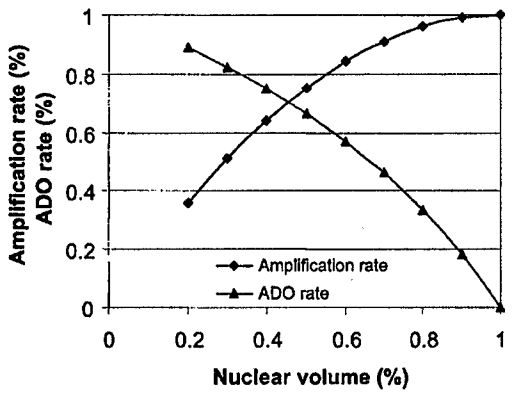


Fig. 3. The effect of the percentage of the total nuclear volume on amplification success and allele dropout. The percentage of total nuclear volume within the tissue section is the X-axis. Amplification success (Y-axis) is the ratio of the number of amplified cells to the total cells expressed as a percentage. Allele dropout (ADO, Y-axis) is the ratio of the number of cells with ADO to the total number of amplified cells expressed as a percentage. If a portion of the nucleus is not present in the section, one or both copies of the *P53* gene may not be present for the amplification. To measure the effect of the percentage of total nuclear volume on amplification and ADO, calculations were performed with the following assumptions: 1) nonuniform distribution of genetic material within the nucleus, and 2) uniform amplification efficiency. For example, if the nuclear volume is 70% of the intact nucleus, the amplification rate is 91% and the ADO rate is 46.2%.

Allele Dropout

ADO is a source of false-negatives. ADO may be caused by failure to retrieve the entire nucleus due to tissue sectioning, degradation or nicking of the genomic DNA template, adsorption of the DNA to the tube surface, preferential amplification of one allele amplification [Liu et al., 1997], or preferential peak generation in the sequencing reaction. Theoretical calculations for spherical nuclei of diameter d uniformly distributed within a tissue section of constant thickness h indicate that the average fraction of nuclear volume is $h/(d+h)$. For an average mammary cell nucleus (4 μm) in a 6-μm section, an average of 60% of the nuclear volume is included. However, this estimate is low, because the microdissected cells were chosen to contain a significant fraction of the nucleus as judged by the intensity of the red stain and the presence of the nucleus on focusing through the section thickness. We estimate that focusing through the microscopic field cannot reliably distinguish a cell with 50% of intact nuclear volume from a cell with 100% of the nuclear volume. Thus, it is possible that an average of 25% of the nuclear volume is not present in the single cells analyzed. Microdissection of partial nuclei is hypothesized to be the major cause of ADO in this situation. If the assumptions of the theoretical calculations are reasonable (see above and Fig. 3 legend), an average nuclear volume of 75% should reduce the amplification rate to 95% and produce an ADO rate of at least 40% (Fig. 3). These

maximum possible values are similar to those observed, compatible with the hypothesis that partial nuclear volume is the major cause of ADO.

False-Positive Immunohistochemical Staining

The P53-stained cells do not necessarily contain a *P53* mutation because the *P53* gene is also activated by hypoxia, heat shock, exposure to nitric oxide, and other stresses. Indeed, the frequency of positive staining depends on the sensitivity of the immunohistochemical staining. With different immunohistochemical staining protocols, a higher frequency of stained cells can be seen and a lower percentage of red cells contain mutations. Double-staining for MDM2 overproduction, which is regulated by P53, was used in an attempt to distinguish physiological induction from the presence of *P53* mutations [Lianes et al., 1994]. However, this protocol was not readily extended from tissue to tissue and this added step did not increase the frequency of observed mutations.

False-Positive Mutations Due to Polymerase Error

Missense mutations in red cells at amino acids conserved in monkey, mouse, rat, and chicken [Soussi et al., 1990] are scored as mutations contributing to mutation load. The false-positive rate was 3% because one missense mutation at an evolutionarily conserved amino acid was observed in 34 clear cells. Four additional, presumably silent sequence changes were also observed (Table II). These false-positive changes may result from a polymerase error in early PCR cycles and the specific mutation is propagated in late cycles to account for ≥20% of the other products.

Mutations are not expected in clear cells. The target size for missense changes at evolutionarily conserved residues is 310 nucleotides. The ratio of missense changes at evolutionarily conserved amino acids (310 nucleotides) to other substitutions (1359 nucleotides) is expected to be 0.23. This is similar to the 1:4 ratio observed in the 34 clear cells. Thus, one might expect about 5.1 artifactual silent changes among the 43 analyzed red mammary cells if the same rate of background occurred. Remarkably, five such changes were identified, compatible with a similar rate of polymerase error. Finally, an analysis of 48 clear cells when SCIMLA was extended to other tissues revealed a false-positive rate of 4%.

The two enzyme system with *Taq* and GB-D DNA polymerases is estimated to have an error rate of 8.5×10^{-6} substitutions per base [Cha and Thilly, 1993]. Once a polymerase error occurs in early PCR cycles, the mutation can be propagated to ≥20% of the other PCR products in late cycles to cause false-positive mutations. This is expected to cause false-positive substitutions in 2.7–5.4% of the amplified cells in the 1.6-kb sequenced region depending on the number of single-stranded templates in single cells.

Caveats

We describe a method of measuring mutation frequency, pattern, and spectrum in situ in normal human tissues. The SCIMLA protocol efficiently allows identification of mutant cells, amplifies a 1.8-kb genome segment, and detects in vivo mutations in single cells from paraffin-embedded, normal human tissues. Caveats include dependence on immunohistochemical staining and thus detection of a subset of primarily missense mutation in the *P53* gene (exons 5–9) that lead to overabundance of P53 protein detectable by immunohistochemical staining. Frameshift mutations (about 10% of total *P53* mutations found in tumors) are missed, although some short in-frame deletions and insertions are detected. Large deletions, retrotransposition events, duplications, and inversions are not detected. The assay involves multiple steps, yet unprecedented measurements are made. Genotyping from all amplified cells is recommended, since a significant number of jackpot mutations occur. Given the jackpot mutations and ADO, amplification and genotyping of 50 or more cells in an individual may be necessary to obtain good estimates of mutation frequency, pattern, and spectrum. Mutation scanning methods such as detection of virtually all mutations-SSCP (DOVAM-S) or denaturing high performance liquid chromatography (DHPLC) can facilitate detection of mutations, especially when many cells must be analyzed due to the presence of jackpot mutations [Oefner and Underhill, 1998; Buzin et al., 2000]. *P53* gene amplification products from 100–150 cells can be genotyped with one set of DOVAM-S gels.

Future Work

SCIMLA may be used to make unprecedented measurements in order to determine the relationship between mutation load and risk for cancer and to evaluate sources of possible mutagen exposure in humans. It may be advantageous to adapt SCIMLA to other organisms, such as rodents and fish. Comparison of mutation load as a function of age and mutagen exposure in transgenic rodent mutation detection systems (e.g., Big Blue) could determine whether the frequency and patterns of *p53* mutations are similar to that previously observed in a bacterial transgene mutation reporter (e.g., the *lacI* gene). Adaptation of SCIMLA to fish could provide a mutagen monitoring system for aquatic environments.

ACKNOWLEDGMENTS

We thank Jie Sheng Chen, Xuemin Li, Martha Magallanes, Jonathan Still, Wenyan Li, and Sofia Loera for technical assistance; William A. Scaringe for help with theoretical calculations for the average fraction of nuclear volume contained within a tissue section; City of Hope Pathology

Core Laboratory for sample acquisition, tissue processing, and immunohistochemical staining; and Rosie Cortez-Rubio for secretarial support. The work was partially supported by DAMD 17-00-1-0204 and R01 AG19784-02.

REFERENCES

Aguilar F, Hussain SP, Cerutti P. 1993. Aflatoxin B_1 induces the transversion of G->T in codon 249 of the p53 tumor suppressor gene in human hepatocytes. Proc Natl Acad Sci USA 90:8586–8590.

Albertini RJ. 2001. HPRT mutations in humans: biomarkers for mechanistic studies. Mutat Res 489:1–16.

Andrew SE, Pownall S, Fox J, Hsiao L, Hambleton J, Penney JE, Kohler SW, Jirik FR. 1996. A novel *lacI* transgenic mutation-detection system and its application to establish baseline mutation frequencies in the *scid* mouse. Mutat Res 357:57–66.

Andrew SE, Reitmair AH, Fox J, Hsiao L, Francis A, McKinnon M, Mak TW, Jirik FR. 1997. Base transitions dominate the mutational spectrum of a transgenic reporter gene in MSH2 deficient mice. Oncogene 15:123–129.

Andrew SE, Xu XS, Baross-Francis A, Narayanan L, Milhausen K, Liskay RM, Jirik FR, Glazer PM. 2000. Mutagenesis in PMS2- and MSH2-deficient mice indicates differential protection from transversions and frameshifts. Carcinogenesis 21:1291–1295.

Bartek J, Iggo R, Gannon J, Lane DP. 1990. Genetic and immunochemical analysis of mutant p53 in human breast cancer cell lines. Oncogene 5:893–899.

Buettner VL, Hill KA, Nishino H, Schaid DJ, Frisk CS, Sommer SS. 1996. Increased mutation frequency and altered spectrum in one of four thymic lymphomas derived from tumor prone p53/Big Blue transgenic mice. Oncogene 13:2407–2413.

Buettner VL, Nishino H, Haavik J, Knöll A, Hill KA, Sommer SS. 1997. Spontaneous mutation frequencies and spectra in p53 (+/+) and p53 (–/–) mice: a test of the "guardian of the genome" hypothesis in the Big Blue transgenic mouse mutation detection system. Mutat Res 379:13–20.

Buzin CH, Wen CY, Nguyen VQ, Nozari G, Mengos A, Li X, Chen JS, Liu Q, Gatti RA, Fujimura FK, Sommer SS. 2000. Scanning by DOVAM-S detects all unique sequence changes in blinded analyses: evidence that the scanning conditions are generic. BioTechniques 28:746–753.

Cariello NF, Skopek TR. 1993. Analysis of mutations occurring at the human hprt locus. J Mol Biol 231:41–57.

Cha RS, Thilly WG. 1993. Specificity, efficiency, and fidelity of PCR. PCR Methods Appl 3:S18–S29.

Cole J, Skopek TR. 1994. Somatic mutant frequency, mutation rates and mutational spectra in the human population in vivo. Mutat Res 304:33–105.

Colgin LMF, Hackmann AFF, Emond MJF, Monnat RJJ. 2002. The unexpected landscape of in vivo somatic mutation in a human epithelial cell lineage. Proc Natl Acad Sci USA JID - 7505876 99:1437–1442.

Dietmaier W, Hartmann A, Wallinger S, Heinmoller E, Kerner T, Endl E, Jauch KW, Hofstadter F, Ruschoff J. 1999. Multiple mutation analyses in single tumor cells with improved whole genome amplification. Am J Pathol 154:83–95.

Ewing B, Green P. 1998. Base-calling of automated sequencer traces using phred. II. Error probabilities. Genome Res 8:186–194.

Ewing B, Hillier L, Wendl MC, Green P. 1998. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. Genome Res 8:175–185.

Gelfand DH, Innis MA, Sninsky JJ. 1999. PCR applications: protocols for functional genomics. In: Gelfand DH, Innis MA, Sninsky JJ, editors. Optimization of PCR: conversations between Michael and David. San Diego: Academic Press. p 3–22.

Hartmann A, Blaszyk H, McGovern RM, Schroeder JJ, Cunningham J, DeVries EMG, Kovach JS, Sommer SS. 1995. p53 gene mutations inside and outside of exons 5-8: the patterns differ in breast and other cancers. Oncogene 10:681–688.

Heddle JA, Dean S, Nohmi T, Boerrigter M, Casciano D, Douglas GR, Glickman BW, Gorelick NJ, Mirsalis JC, Martus HJ, Skopek TR, Thybaud V, Tindall KR, Yajima N. 2000. In vivo transgenic mutation assays. Environ Mol Mutagen 35:253–259.

Heinmoller E, Liu Q, Sun Y, Schlake G, Hill KA, Weiss LM, Sommer SS. 2002a. Toward efficient analysis of mutations in single cells from ethanol-fixed, paraffin-embedded, and immunohistochemically stained tissues. Lab Invest 82:443–453.

Heinmoller E, Schlake G, Renke B, Liu Q, Hill KA, Sommer SS, Ruschoff J. 2002b. Microdissection and molecular analysis of single cells or small cell clusters in pathology and diagnosis — significance and challenges. Anal Cell Pathol (in press).

Hill KA, Buettner VL, Glickman BW, Sommer SS. 1999. Spontaneous mutations in the Big Blue transgenic system are primarily mouse derived. Mutat Res Minirev 436:11–19.

Hussain SP, Raja K, Amstad PA, Sawyer M, Trudel LJ, Wogan GN, Hofseth LJ, Shields PG, Billiar TR, Trautwein C, Hohler T, Galle PR, Phillips DH, Markin R, Marrogi AJ, Harris CC. 2000. Increased p53 mutation load in nontumorous human liver of Wilson disease and hemochromatosis: oxyradical overload diseases. Proc Natl Acad Sci USA 97:12770–12775.

Hussain SP, Amstad P, Raja K, Sawyer M, Hofseth L, Shields PG, Hewer A, Phillips DH, Ryberg D, Haugen A, Harris CC. 2001. Mutability of p53 hotspot codons to benzo(a)pyrene diol epoxide (BPDE) and the frequency of p53 mutations in nontumorous human lung. Cancer Res 61:6350–6355.

Jonason AS, Kunala S, Price GJ, Restifo RJ, Spinelli HM, Persing JA, Leffell DJ, Tarone RE, Brash DE. 1996. Frequent clones of p53-mutated keratinocytes in normal human skin. Proc Natl Acad Sci USA 93:14025–14029.

Lewis PD, Harvey JS, Waters EM, Skibinski DO, Parry JM. 2001. Spontaneous mutation spectra in supF: comparative analysis of mammalian cell line base substitution spectra. Mutagenesis 16:503–515.

Lianes P, Orlow I, Zhang ZF, Oliva MR, Sarkis AS, Reuter VE, Cordon-Cardo C. 1994. Altered patterns of MDM2 and TP53 expression in human bladder cancer. J Natl Cancer Inst 86:1325–1330.

Liu Q, Thorland EC, Sommer SS. 1997. Inhibition of PCR amplification by a point mutation downstream of a primer. BioTechniques 22:292–294, 296, 298, passim.

Liu W, Li X, Chen JS, Sommer SS. 2003. Robust dosage–PCR for detection of heterozygous chromosomal deletions. Bio Techniques 34:558–568.

Mitnik L, Heller C, Prost J, Viovy JL. 1995. Segregation in DNA solutions induced by electric fields. Science 267:219–222.

Motorna OO, Martin H, Gentile GJ, Gentile JM. 2001. Analysis of lacI mutations in Big Blue transgenic mice subjected to parasite-induced inflammation. Mutat Res 484:69–76.

Nohmi T, Suzuki T, Masumura K. 2000. Recent advances in the protocols of transgenic mouse mutation assays [In Process Citation]. Mutat Res 455:191–215.

Oefner P, Underhill P. 1998. DNA mutation detection using denaturing high-performance liquid chromatography (DHPLC). Curr Protoc Hum Genet Suppl 19:7.10.1–7.10.12.

Ono T, Ikehata H, Nakamura S, Saito Y, Hosoi Y, Takai Y, Yamada S, Onodera J, Yamamoto K. 2000. Age-associated increase of spontaneous mutant frequency and molecular nature of mutation in newborn and old lacZ-transgenic mouse. Mutat Res 447:165–177.

Recio L, Cochrane J, Simpson D, Skopek TR, O'Neill JP, Nicklas JA,

Albertini RJ. 1990. DNA sequence analysis of in vivo hprt mutation in human T lymphocytes. Mutagenesis 5:505–510.

Sands AT, Suraokar MB, Sanchez A, Marth JE, Donehower LA, Bradley A. 1995. p53 deficiency does not affect the accumulation of point mutations in a transgene target. Proc Natl Acad Sci USA 92:8517–8521.

Skopek TR, Kort KL, Marino DR. 1995. Relative sensitivity of the endogenous hprt gene and lacI transgene in ENU-treated Big Blue B6C3F1 mice. Environ Mol Mutagen 26:9–15.

Skopek TR, Kort KL, Marino DR, Mittal LV, Umbenhauer DR, Laws GM, Adams SP. 1996. Mutagenic response of the endogenous *hprt* gene and *lacI* transgene in benzo[a]pyrene-treated Big Blue B6C3F1 mice. Environ Mol Mutagen 28:376–384.

Soussi T, Caron de Fromentel C, May P. 1990. Structural aspects of the p53 protein in relation to gene evolution. Oncogene 5:945–952.

Stuart GR, Yoshimitsu O, de Boer JG, Glickman BW. 2000. Mutation frequency and specificity with age in liver, bladder and brain of *lacI* transgenic mice. Genetics 154:1291–1300.

Vogelstein B, Kinzler KW. 1999. Digital PCR. Proc Natl Acad Sci USA 96:9236–9241.

Walker VE, Gorelick NJ, Andrews JL, Craft TR, de Boer JG, Glickman BW, Skopek TR. 1996. Frequency and spectrum of ethylnitrosourea-induced mutation at the *hprt* and *lacI* loci in splenic lymphocytes of exposed *lacI* transgenic mice. Cancer Res 56:4654–4661.

Walker DR, Bond JP, Tarone RE, Harris CC, Makalowski W, Boguski MS, Greenblatt MS. 1999. Evolutionary conservation and somatic mutation hotspot maps of p53: correlation with p53 protein structural and functional features. Oncogene 18:211–218.

Wetmur JG. 1991. DNA probes: applications of the principles of nucleic acid hybridization. Crit Rev Biochem Mol Biol 26:227–259.

Accepted by—
W.H. Koch

0023-6837/02/8204-443$03.00/0
LABORATORY INVESTIGATION
Copyright © 2002 by The United States and Canadian Academy of Pathology, Inc.

APPENDIX 2

Vol. 82, No. 4, p. 443, 2002
Printed in U.S.A.

Toward Efficient Analysis of Mutations in Single Cells from Ethanol-Fixed, Paraffin-Embedded, and Immunohistochemically Stained Tissues

Ernst Heinmöller, Qiang Liu, Yuan Sun, Gudrun Schlake, Kathleen A. Hill, Larry M. Weiss, and Steve S. Sommer

Department of Molecular Genetics (EH, QL, GS, KAH, SSS), City of Hope National Medical Center/Beckman Research Center; and Department of Pathology (YS, LMW), City of Hope National Medical Center, Duarte, California

SUMMARY: Only a few studies have demonstrated successful molecular analysis after whole genome amplification using single cells dissected from paraffin-embedded tissues. The results in these studies were limited by low-amplification efficiency and high rates of allele dropout. In the present study, the amplification rate using a thoroughly modified primer extension and preamplification-PCR protocol was improved significantly for single cells microdissected from paraffin-embedded and immunohistochemically stained tissues. Tissue fixation with ethanol (85%) and the addition of 0.2 mmol/L EDTA helped to achieve an amplification rate between 67% (segments 200 to 400 bp) and 72% (segments <200 bp). Normal tissue sections were immunohistochemically double stained for overabundance of p53 protein and proliferating cell nuclear antigen. Microdissection of single cells was performed with a manual micromanipulator equipped with a Tungsten needle. Sequence analysis of the TP53 gene was performed after improved primer extension preamplification-PCR and multiplex PCR from single microdissected cells. The rate of allele dropout was at least 68%. These technical advances facilitate routine mutation analysis using a single cell or a few cells microdissected from routinely processed paraffin-embedded normal and tumor tissues. Allele dropout still represents a serious problem in single-cell mutation analysis, especially in samples with limited template DNA and prone to DNA damage. (*Lab Invest 2002, 82:443–453*).

Molecular analysis of single cells has become routine in reproductive medicine (Hahn et al, 2000). However, few studies have routinely analyzed rare heterozygous mutations in single cells microdissected from fixed, paraffin-embedded, and immunohistochemically stained tissue sections because of certain limitations (Becker et al, 1996; Roehrl et al, 1997; Schutze and Lahr, 1998; Suarez-Quian et al, 1999). Crosslinking of DNA in tissues fixed with the commonly used fixative, formalin, limits the efficiency of PCR amplification and reduces the maximum length of the PCR amplicon (Karlsen et al, 1994; Williams et al, 1999). Allele dropout (ADO) occurs when one allele is preferentially amplified over the other to the extent that the minor allele is not detected. ADO is particularly problematic when templates consisting of DNA from less than 10 cells are amplified; it occurs at the same rate irrespective of whether nested PCR or whole genome amplification (WGA) is used for initial amplification (Hahn et al, 1998, 2000). ADO is especially problematic in the diagnosis of rare heterozygous mutations from single cells. No protocol has been identified that eliminates ADO (Hahn et al, 2000). Of great interest is the development of methodologies that permit examination of rare heterozygous mutations in single cells from normal tissues with high sensitivity and accuracy to enable investigation of the role of mutagenesis and risk of disease.

Immunohistochemistry is a technique that specifically provides information about protein expression patterns at the single-cell level and may indicate molecular changes relevant to neoplastic development. The tumor suppressor gene *TP53* is mutated in 50% of human cancers (Greenblatt et al, 1994). In its mutated form or under certain cellular and/or genotoxic stress, the p53 protein becomes stabilized and its overabundance is readily detectable by immunohistochemical staining (May and May, 1999; Wallace-Brodeur and Lowe, 1999). Overabundance of p53 protein may be found even in normal tissues in the nuclei of cells with a morphologically normal appearance (Barnes et al, 1992; Huusko et al, 1999; Varley et al, 1999). Immunohistochemical staining for p53 can be used to identify cells that are potential candidates for *TP53* gene mutations. Because proliferating cell nuclear antigen (PCNA) expression is regulated by p53, double immunostaining for overexpression of p53 *and* PCNA proteins was used to enhance detection of cells with overexpression of p53 caused by a mutation in the p53 gene of functional relevance.

Received November 8, 2001.
Supported in part by the Deutsche Krebshilfe (to EH). *Current address for Dr. Heinmöller: Institute of Pathology, Klinikum Kassel, Mönchebergstrasse 41-42, 34125 Kassel, Germany.*
Address reprint requests to: Dr. Steve S. Sommer, Chair, Department of Molecular Genetics, Beckman Research Institute/City of Hope, 1450 East Duarte Road, Duarte, CA 91010-0269. E-mail: sommerlab@coh.org

In general, for molecular analyses from single cells, two different PCR techniques can be used: direct (usually nested) amplification or (nonspecific) preamplification of the genetic material by WGA followed by specific (nested) PCR. A major advantage of WGA is the possibility of multiple genetic analyses from a single cell. Of the various WGA techniques available, primer extension and preamplification (PEP)-PCR and degenerate oligonucleotide-primed PCR provide the most complete coverage of the genome (Wells et al, 1999). Improved primer extension and preamplification (I-PEP)-PCR is superior to degenerate oligonucleotide-primed PCR in amplification efficiency (Dietmaier et al, 1999). PEP-PCR is a technique of WGA used before specific PCR from minute amounts of DNA, which enables the molecular analysis of multiple genes from a single cell or a few cells (Zhang et al, 1992). PEP-PCR has recently been improved to facilitate multiple genetic analyses of small cell clusters microdissected from formalin-fixed paraffin-embedded tissues (Dietmaier et al, 1999). Although some studies have analyzed DNA from immunohistochemically stained single cells from fresh-frozen tissue samples (Persson et al, 2000; Ponten et al, 1997), to our knowledge, no study has conducted WGA before specific amplification of single-cell genomes from fixed, paraffin-embedded, and immunohistochemically stained tissue sections. Herein, DNA integrity, amplification efficiency, and ADO are evaluated after an overall enhanced protocol for microdissection of single cells from ethanol-fixed, paraffin-embedded, and immunohistochemically stained tissue sections and amplification with a modified I-PEP-PCR protocol.

In this study major improvements to tissue fixation and processing and single-cell DNA amplification were made to enable the analysis of rare heterozygous mutations in the *TP53* gene of single cells microdissected from ethanol-fixed, paraffin-embedded, and immunohistochemically stained sections of normal human tissue (Fig. 1).

Results

Tissue Fixation and Immunohistochemistry

To avoid DNA crosslinking by formalin fixation of tissue, which limits molecular analysis of single cells,

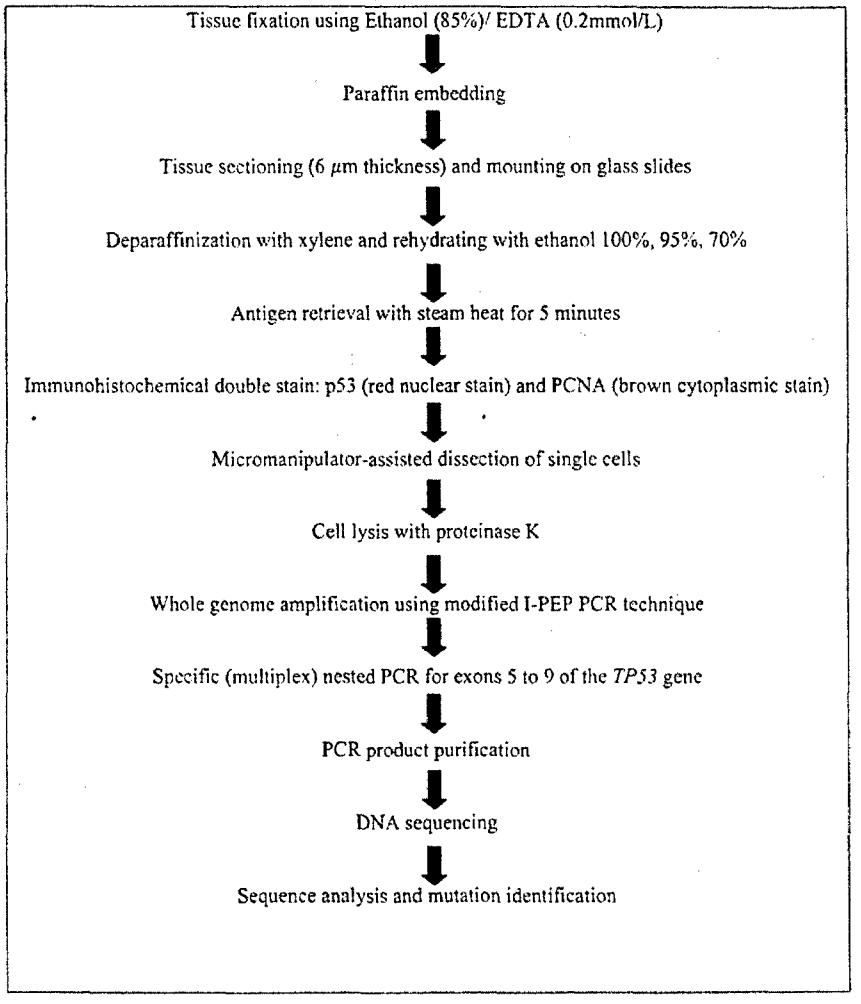


Figure 1.
A flow chart of the methods used for mutation analysis using ethanol-fixed, paraffin-embedded, and immunohistochemically stained single cells or small cell clusters

ethanol/EDTA fixation was optimized to maintain high DNA integrity and provide high-resolution immunohistochemistry. Ethanol has a precipitating effect on tissues. EDTA inhibits DNases present in surgically removed tissues for better preservation of DNA. Proteinase K digestion of various tissues fixed with ethanol/EDTA ("Materials and Methods") revealed the presence of single-stranded, high molecular weight DNA (~20 Kb) after the final step of immunohistochemical staining in all tissues analyzed (data not shown). To demonstrate the feasibility of ethanol/EDTA tissue fixation for immunohistochemistry, double staining of p53 protein (red-colored nuclear stain) together with PCNA (brown-colored nuclear stain) was used. A mix of 85% ethanol/0.2 mmol/L EDTA was found to give crisp and clear immunohistochemistry after a brief antigen retrieval by steam treatment at 95° C for 5 minutes, enabling the clear distinction of nuclear staining for p53 protein, or PCNA, or both p53 and PCNA in various tissues with physiologically different expression patterns of both proteins (Fig. 2). Fixation with 95% ethanol/0.1 mmol/L EDTA or 75% ethanol/1 mmol/L ETDA resulted in less clear staining (not shown). In addition, cutting of sections was impaired when 95% ethanol/0.1 mmol/L EDTA was used as fixative.

Optimization of I-PEP-PCR: Analysis of Dilutions of Genomic DNA

Preliminary experiments indicated that the amplification success rate of various exons of the TP53 gene from single cells microdissected from ethanol/EDTA-fixed tissues was 30% or less when the I-PEP technique originally described by Dietmaier et al (1999) was used. Such a low amplification success rate from single cells would make the routine analysis of genes like TP53 costly and cumbersome, because exons 5 to 9, which are commonly examined, would be amplified in only a small number of microdissected single cells. To optimize I-PEP, genomic DNA was serially diluted to the single-cell level (6 pg). To ensure optimal digestion, the concentration of Tween-20 in the digestion buffer was increased from 0.5% to 3%. To achieve maximum amplification efficiency, various PCR cycling conditions (annealing temperature, extension time) and I-PEP-PCR mix ingredients (concentration of primers, $MgCl_2$, dNTP, enzymes) were tested. The concentrations of $MgCl_2$ and dNTP were increased stepwise from 2.5 mmol/L and 0.1 mmol/L to 6.0 mmol/L and 1 mmol/L, respectively. Specific PCR after I-PEP-PCR was most effective when $MgCl_2$ was 6.0 mmol/L and dNTPs were 1 mmol/L each in I-PEP-PCR. Of the primer concentrations tested (5 μmol/L, 10 μmol/L, 15 μmol/L, 20 μmol/L, 25 μmol/L, 30 μmol/L), 20 μmol/L was found to be optimal. Of the various enzymes tested (*Taq* Expand High Fidelity Polymerase with and without Bst DNA polymerase [Roche, Basel, Switzerland], Pfu Turbo [Stratagene, La Jolla, California], and AdvanTaq [Clontech Laboratories, Palo Alto, California]), *Taq* Expand High Fidelity Polymerase without Bst DNA polymerase was found to be superior. No amplification was seen when Pfu Turbo was used in I-PEP (not shown). Lowering the annealing temperature from 37° C to 28° C, reducing the annealing time from 4 minutes to 2 minutes, and prolonging the extension time from 30 seconds to 3 minutes enabled the amplification of exons 5 to 9 of

Figure 2.
Immunohistochemical staining for p53 and PCNA protein (overabundance) in nucleus and cytoplasm, respectively, for normal human tissues that have been fixed in ethanol and embedded in paraffin. A, Normal colon tissue with crypt epithelial cells staining positive for PCNA (brown cytoplasmic stain) and a single cell staining positive (red nuclear stain) for p53 overabundance (*arrow*) (magnification ×100). B, Normal lung tissue with bronchial epithelial cells. A cell staining for PCNA (brown cytoplasmic stain, *white arrow*) can be seen in addition to a cell staining for p53 overabundance (red nuclear stain, *black arrow*) (magnification ×40). C, Benign proliferative lesion of the mammary gland. A cell staining red for p53 protein overabundance (*white arrow*) is seen together with a cell staining mixed red-brown for both p53 and PCNA (*black arrow*) proteins and many other cells staining brown for PCNA overabundance (magnification ×250).

the *TP53* gene after I-PEP-PCR. A nested amplification of ~2 Kb from high molecular weight genomic DNA was possible even to a dilution of 13.2 pg of template DNA, which is equivalent to four copies of mammalian genome.

Measurement of ADO After Amplification from Single Human Fibroblasts

Amplification efficiency using single fibroblasts (HF57 human fibroblasts) was assessed using amplification of various segment lengths within exons 5 to 9 of the *TP53* gene in a nested PCR configuration after I-PEP. Unfixed fibroblasts were air dried for 2 hours and used for analysis of amplification from single cells and for measurement of ADO. Segments of ~500 bp (eg, spanning exons 5 and 6) were amplified with a success rate of 68% (33 of 48 single cells); segments of 300 bp or less (eg, exon 6) were amplified with a success rate of 87% (42 of 48 single cells; not shown). A ~2 Kb segment (spanning exons 5 to 9) could be amplified in 4 (8%) of 50 single cells after I-PEP.

HF57 human fibroblasts are heterozygous for a polymorphism (G>A bp 13494) in intron 6 of the *TP53* gene, which makes them useful for the study of ADO after I-PEP-PCR from single cells. Exon 6 was amplified after I-PEP-PCR from 37 air-dried, single HF57 fibroblast cells. ADO was determined after DNA sequencing. The ADO rate was 68% (25 of 37 single cells). Twelve single cells (32%) demonstrated heterozygosity at bp 13494. Of the single cells with ADO, 32% (12 single cells) showed the mutant allele sequence G>A bp 13494, and 35% (13 single cells) exhibited the wild-type sequence (Fig. 3). ADO was not observed when five or more cells were used for I-PEP-PCR.

Amplification from Single Cells Microdissected from Ethanol-Fixed, Paraffin-Embedded, and Immunohistochemically Stained Tissue Sections

Amplification efficiency using single microdissected cells from ethanol-fixed, paraffin-embedded, and immunohistochemically stained tissue sections was determined by amplification of various exons of the *TP53* gene after I-PEP-PCR and a nested PCR configuration. First-round amplification was performed in a multiplex PCR using primers covering exons 5 to 9 of the *TP53* gene. Segments between 200 and 400 bp (single exons of *TP53* gene) were amplified in up to 67% of cells; segments smaller than 200 bp were amplified in up to 72% of cells (Fig. 4). Amplification efficiency was dependent on the type of tissue analyzed. Pancreatic tissue has a high degree of enzymatic DNA degradation starting early during surgical manipulation. DNA amplification from pancreatic tissue was compared with colon tissue, in which DNA is usually not affected by enzymatic damage during surgery. As expected, amplification success was greater using single cells from colon than from pancreatic tissue. In a series of amplification experiments, all five exons (5 to 9) of the *TP53* gene were amplified in 50% of single cells harvested from colon versus

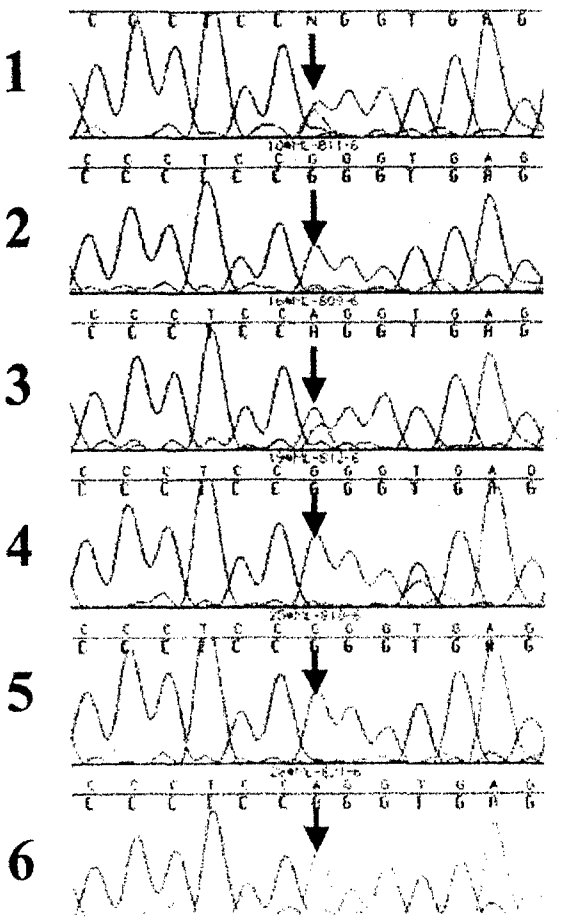


Figure 3.
Evaluation of allele dropout (ADO) after improved primer extension and preamplification (I-PEP)-PCR. Sequence analysis of a heterozygote sequence change intron 6 bp 13494 (G>A) of TP53 gene in single HF57 fibroblasts. *Lanes 1 and 3*: single cell, heterozygosity; *lanes 2, 4, and 5*: single cell, wild-type, allele dropout; *lane 6*: single cell, G>A, ADO.

36% from pancreatic tissue. The amplification rates of at least two exons were considerably higher in colon tissues (Table 3). High multiplex amplification success rate was achieved using tagged primers (see "Materials and Method") in first-round PCR in addition to an antibody that ensures a hot start PCR by neutralizing *Taq* polymerase.

Identification of Rare Heterozygous Mutations in Single Cells of Normal Colon Identified To Have an Overabundance of P53 Protein

Normal colon tissue from five patients who were treated for colon carcinoma was fixed in ethanol/EDTA, embedded in paraffin, and immunohistochemically stained for identification of p53 protein overexpression. Normal tissues from two of these patients did not stain positive for p53 protein. Normal tissues from the remaining three patients exhibited a positive nuclear staining for p53 protein in up to nine epithelial cells/slide; the total number of epithelial cells per slide was estimated to be between 40,000 and 60,000 cells. A total of 23 single cells and a cell cluster (40 cells)

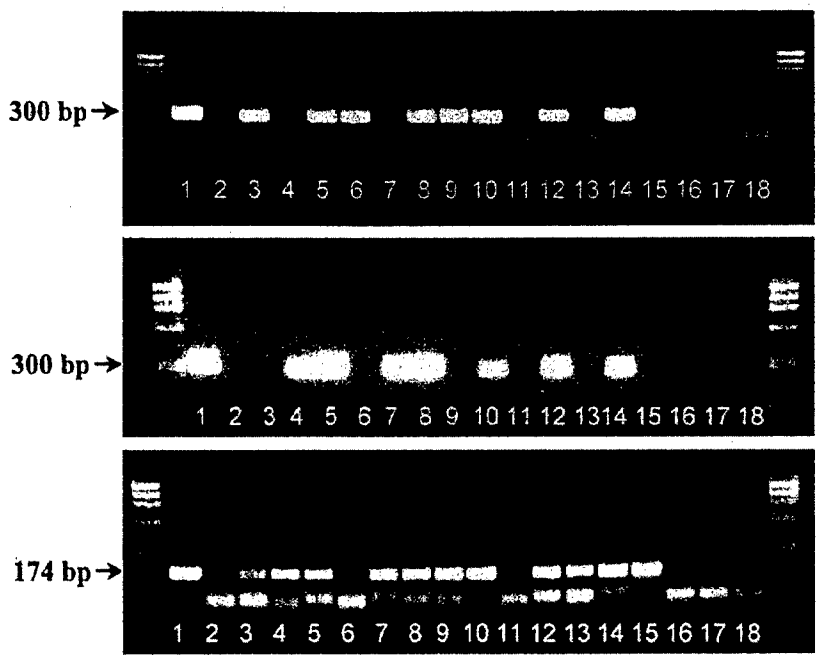


Figure 4.
Representative examples of the amplification rate of various segments of exons 5 to 9 from colon cells after I-PEP-PCR. Upper: Nested PCR exon 9 (first-round segment 378 bp); middle: different samples with nested PCR exon 5 (first-round segment 383 bp); lower: different samples with nested PCR exon 5 (first-round segment 181 bp). *Lane 1*: 30 cells; *lanes 2 to 16*: single cell; *lanes 17 and 18*: negative control, no cells.

with positive p53 protein immunohistochemistry were microdissected and amplified. Sequence analysis revealed five sequence changes (Fig. 5).

Discussion

Formalin Fixation Limits Molecular Analysis, Especially Using Single Cells

There is no better place to study the underlying molecular genetic dynamics of normal and pathologic processes than in vivo. Histologic tissue sections are one way to gain access to a fingerprint of these processes within a defined morphologic compartment and at the level of a single cell (Emmert-Buck et al, 1996; Schutze and Lahr, 1998). However, technical limitations have hindered widespread genomic analyses especially at the single-cell level in large part because of the routine formalin fixation of tissues, which causes DNA crosslinking, thereby reducing PCR efficiency. This is one of the reasons why the achieved amplification rates reported from single cells microdissected from formalin-fixed paraffin-embedded tissue slides did not exceed 35% in short segments up to 246 bp (Becker et al, 1996; Roehrl et al, 1997). Considerably higher amplification rates (up to 70%) have been reported from single cells dissected from frozen, immunohistochemically stained tissues (Persson et al, 2000). However, major limitations in the analysis of frozen tissue sections include the usually considerably lower quality of immunohistochemical staining, especially if double staining is performed. In addition, the availability of frozen tissues for molecular analyses is usually much lower in comparison to paraffin-embedded tissues. To circumvent the disadvantages associated with formalin fixation and frozen tissues, tissues were fixed in an ethanol/EDTA solution before paraffin embedding.

Ethanol/EDTA Fixation Improves Molecular Analyses Using Single Cells

In the present study, a mixture of 85% ethanol/0.2 mmol/L EDTA gave crisp and clear immunohistochemical staining for two proteins (p53 and PCNA) simultaneously in various tissues, providing the basis for coupled morphologic and molecular analyses in paraffin-embedded tissues at a single cell-level. It has been demonstrated that ethanol fixation is the method of choice for immunohistochemical staining of a variety of tissue antigens, especially the p53 protein (Arnold et al, 1996; Bassarova and Popov, 1998). Furthermore, addition of EDTA has been shown to inhibit DNA degradation in formalin-fixed tissues to a certain degree (Yagi et al, 1996). The specificity of the immunohistochemical double staining for both proteins was demonstrated by frequent expression of both proteins in highly dividing tissues like colon and a proliferative lesion of the mammary gland, by moderate expression of both proteins in moderately dividing tissues like bronchial epithelium, and by rare expression of one of the proteins in slowly dividing tissues like liver tissue.

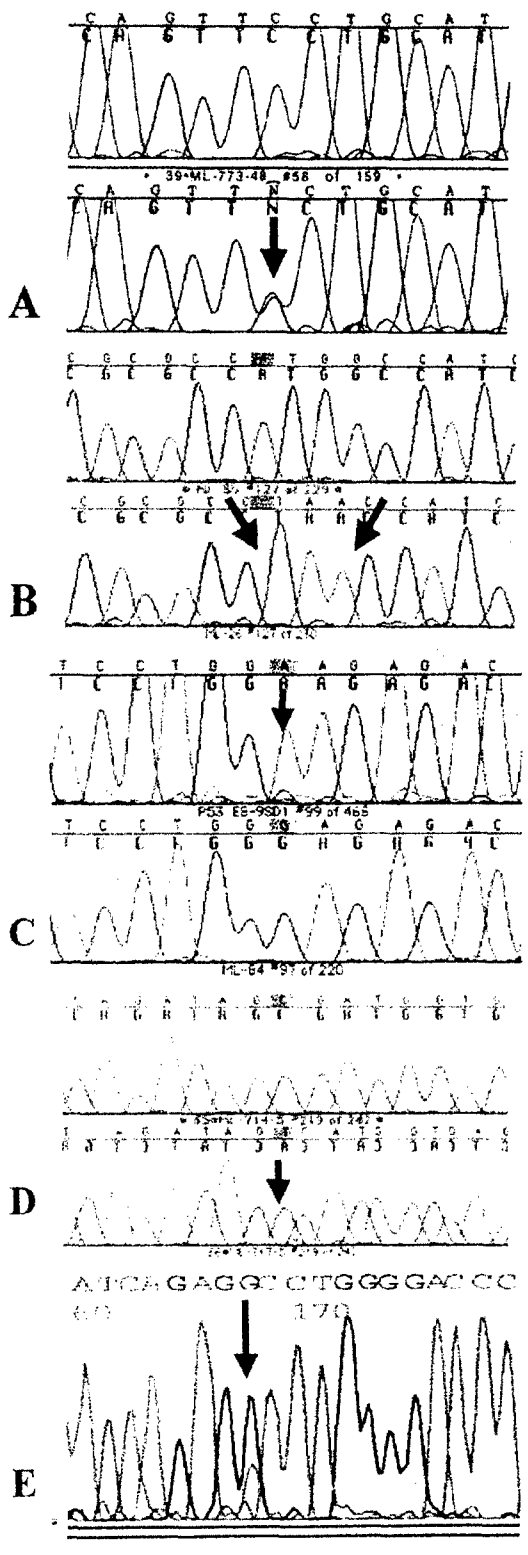


Figure 5.
Identification of mutations in exons 5 to 9 of the *TP53* gene from single cells that were immunohistochemically stained positive for p53 protein overabundance. Please note that *arrows* are used to mark the mutant sequence. The single cells were from paraffin-embedded, normal colon tissue. A, Forty cells with a heterozygous mutation: C to T transition (*arrow*) exon 7 bp 14,049 (pro/val). B, Single cell with deletion of four bases (ATGG) and insertion of three bases ("TAA" arrows) exon 5 (bp 13,157 to 13,161). C, Single cell with a silent G to A transition (*arrow*) in exon 8 (bp 14,506). D, Single cell with a C to T transition (*arrow*) in exon 5 (bp 13,234, asp/val). E, Single cell with a heterozygous sequence change in intron 5 (G to A transition, bp 13,292).

Modified I-PEP Shows Improved Amplification Success

Herein, a recently published I-PEP-PCR technique (Dietmaier et al, 1999) was thoroughly optimized. The revised protocol demonstrated amplification of ~2 Kb segments after I-PEP-PCR when high molecular weight DNA was used as the template, even using dilutions down to 12 pg, ie, four copies (or two cells) of mammalian genome. Successful amplification of 2 Kb segments after PEP has been described previously but only after using 5000 cells as initial template (Duddy et al, 1998). Using air-dried single fibroblasts, a 2 Kb segment could be amplified but only with a small number of cells. This may be a result of the use of unfixed cells that were air dried for at least 2 hours, when oxygen-related damage may have already led to DNA degradation. When smaller segments were amplified after I-PEP-PCR, amplification efficiency was up to 87%. Most importantly, the amplification rate of segments covering single exons of the *TP53* gene was increased from 30% to up to 72% from single ethanol-fixed, paraffin-embedded, and immunohistochemically stained single cells after the present modifications of I-PEP-PCR. As a result of this improvement, serial molecular analyses of single cells stained for overabundance of proteins like p53 or others can be done in a reasonable number of PCR experiments. Exons 5 to 9 of the *TP53* gene were amplified in 50% of single cells dissected from colon samples. The amplification success rate in single pancreatic cells was lower, most likely because of the much higher degree of pancreatic protease-related DNA damage (autodigestion) beginning early during surgical manipulation of this tissue, in addition to the long surgical procedure if performed as Whipple procedure. But even under these most unfavorable conditions, multiplex PCR after WGA offers the opportunity of multiple genetic analyses of various genes even if multiple exons are to be analyzed. In addition, if mutations are found by sequence analysis, reamplification for confirmation of a particular base change can be performed from the original template, which is not possible when single cells are amplified directly by specific PCR without prior preamplification by any WGA method. Of interest is the rate of polymerase errors during the amplification from single cells. Herein, over 10,000 nucleotides were sequenced in the analysis of single, air-dried fibroblasts and not a single artificial mutation was observed.

ADO After Modified I-PEP Amplification from Single Cells Is Still High

ADO or preferential amplification of one allele over the other during PCR amplification is a well-known phe-

Table 1. Protocols for Cell Digestion and PEP-PCR Mixes

A. This study	B. Dietmaier et al. (Am J Pathol 1999;154:83–95)
• Tween-20: 3%	Tween-20: 0.5%
• Proteinase K: 0.5 mg/ml	Proteinase K: 4 mg/ml
• PEP-primer: 20 μmol/L	Pep-primer: 16 μmol/L
• $MgCl_2$: 6 mmol/L	$MgCl_2$: 2.5 mmol/L
• Gelatin: 0.05 mg/ml	Gelatin: 0.05 mg/ml
• dNTPs: 4 mmol/L	dNTPs: 0.1 mmol/L
• DMSO: 5%	DMSO: 5%
• Taq Expand HiFi: 2.5 U	Taq Expand HiFi: 1.8 U

Final concentrations are given.

nomenon limiting detection of heterozygote base changes (Hahn et al, 1998, 2000). ADO rates between 0% and 83% have been reported and are dependent on the source and quality of the template DNA amplified (Findlay et al, 1995, 1998; Garvin et al, 1998; Hahn et al, 1998; Rechitsky et al, 1996; Snabes et al, 1994; Wells et al, 1999). As a result of ADO, mutation analysis may fail to detect heterozygous mutations in a significant number of cases. The distribution of alleles dropping out during PCR amplification occurs at random. The occurrence of ADO is influenced by multiple factors, eg, cell lysis, conditions of PCR mix, thermal cycling conditions, guanine/cytosine content of the DNA segment amplified, or degradation of the DNA template (El-Hashemite and Delhanty, 1997; Ray and Handyside, 1996; Walsh et al, 1992). Furthermore, ADO occurs in either direct, specific (nested) amplification or in PEP-PCR before specific PCR from single cells (Hahn et al, 1998). This ADO estimate is specifically for the amplification method and is a minimal estimate because it is expected that ADO is much higher for analysis of single cells microdissected from paraffin-embedded tissues. No method described to date has been able to eliminate ADO.

Forty-one cell culture human fibroblasts harboring a heterozygous polymorphism in intron 6 of the *TP53* gene were used for the analysis of ADO using the modified I-PEP-PCR protocol. Fibroblasts were spread on slides, air dried, and stained with methylene blue. Because fibroblasts were not prepared with a fixative like ethanol or acetone to ensure at least a limited protection, the fibroblasts were not protected from oxygen radical–induced DNA damage. Because DNA degradation starts immediately after cells are exposed to oxygen, this situation closely resembles the exposure of immunohistochemically stained cells to DNA damaging oxygen radicals while sitting as a single-cell layer on slides in the course of tissue

Table 2. PCR Primers

PCR primers	5′ to 3′ Sequence
First round	
E5-9PD2:	5′ gcggtcccaaaagggtcagtctcgctagtgggttgcagga 3′
E5PU1X:	5′ gcggtcccaaaagggtcagtaagagcaatcagtgaggaa tc 3′
E6PD1X:	5′ gcggtcccaaaagggtcagtagctggggctggagagac 3′
E6PU1X:	5′ gcggtcccaaaagggtcagtcaaataagcagcaggagaaag 3′
E7PD1X:	5′ gcggtcccaaaagggtcagtccctgcttgccacaggtc 3′
E7PU2X:	5′ gcggtcccaaaagggtcagtaagaaatcggtaagaggtggg 3′
E8PD2X:	5′ gcggtcccaaaagggtcagtttgggagtagatggagcctg 3′
E8PU1X:	5′ gcggtcccaaaagggtcagtgaaagaggcaaggaaaggtg 3′
E9PD1X:	5′ gcggtcccaaaagggtcagtaagcaggacaagaagcggtg 3′
E8-9SU1X:	5′ gcggtcccaaaagggtcagtggagccattgtctttgaggc 3′
Second round	
E5-26D:	5′ ttcacttgtgccctgactt 3′
E5-291U:	5′ ccgttggtcgggacagc 3′
E6SD2:	5′ acagggctggttgcccag 3′
E6SU2:	5′ gccactgacaaccaccctt 3′
E7SD2:	5′ ccaaggcgcactggcctc 3′
E7SU1:	5′ tggggcacagcaggccag 3′
E8PD3:	5′ ttccttactgcctcttgcttc 3′
E8PU2:	5′ agtgaatctgaggcataactg 3′
E9PD2:	5′ gagaccaagggtgcagttatg 3′
E9PU1:	5′ ggcaaatgccccaattgcag 3′

U = Upstream; D = Downstream; X = (5′ tagged 20-mer universal sequence); P = oligo designed for PCR amplification; S = oligo was designed for DNA sequencing.

Table 3. Amplification Efficiency of Exons 5 to 9 of the *TP53* Gene[a]

Number of exons	Amplification	%
Colon tissue		
2 (e.g. exons 6+8)[b]	17/18	94
3	16/18	89
4	12/18	66
5 (exons 5–9)	9/18	50
Pancreatic tissue		
2 (e.g. exons 6+8)	17/28	61
3	15/28	54
4	12/28	43
5 (exons 5–9)	10/28	36

[a] Eighteen microdissected paraffin-embedded single cells from normal colon tissue and 28 single cells from normal pancreatic tissue.
[b] At least two exons.

processing and microdissection. Herein the ADO rate was 68% and affected both alleles randomly. If a heterozygous mutation is present, according to the data herein, the chance of the correct diagnosis of heterozygosity will be 22%.

Mutation Analysis in Single Cells from Ethanol-Fixed, Paraffin-Embedded, and Immunohistochemically Stained Normal Tissues

This is the first report of specific molecular analyses after WGA from immunohistochemically stained single cells microdissected from paraffin-embedded tissues. In the analysis of a limited number of microdissected single cells from normal colon tissue staining positive for p53 protein, sequence changes in the *TP53* gene were found in 4 of 23 single cells and in 1 cell cluster of 40 cells. The sequence changes were not published mutational hotspots of the *TP53* gene. The relevance of the heterozygote mutations detected in intron 5 of one single cell and the heterozygous base change seen in the cell cluster in regard to inactivation of the *TP53* gene is not clear. However, heterozygous mutations inactivating the remaining wild-type allele have been described (Forrester et al, 1995; Hachiya et al, 1994; Milner and Medcalf, 1991; Villadsen et al, 2000). Possible reasons besides ADO for detecting only wild-type sequence in a subset of single cells with positive p53 immunohistochemistry dissected from tissue slides may be contamination of the PCR or loss of genetic material containing the allele harboring the base change during tissue cutting. The chance of losing genetic material from a single cell nucleus in the course of tissue sectioning can be estimated to be between 10% and 20% if sections of 6 μm are processed. It should be noted that switching from the analysis of single paraffin-embedded cells to single cells from fresh-frozen tissues might not be helpful in solving the ADO problem. A recently published article reported an ADO rate of 50% from single cells dissected from immunohistochemically stained fresh-frozen tissue sections (Persson et al, 2000).

Conclusion

In summary, we present a reproducible and reliable fixation protocol for paraffin-embedded tissues that allows clear and crisp immunohistochemical staining of at least two proteins simultaneously and preserves DNA integrity to a degree that allows the molecular analysis of multiple genes at the single-cell level. Single-cell microdissection was achieved by a fast and easy-to-handle simple manual microdissection device that is affordable to many laboratories, being about one tenth the price of a laser microdissection device. Furthermore, the success rate of PCR amplification from DNA obtained from a single cell dissected from paraffin-embedded tissue is increased to a level where serial analyses of mutations in a single cell or a few cells from normal or tumor tissue can be performed in a reasonable number of experiments. This could be the basis for further studies to determine an individual *TP53* mutation load in normal tissues in the context of morphology, enabling an estimation of individual cancer risk well in advance of the development of malignant disease.

Material and Methods

Cell and Tissue Samples

Tissue samples from patients in surgery for colorectal cancer or pancreatic cancer were obtained from the Department of Pathology from the City of Hope National Medical Center (Duarte, California). Suspensions of cultured human fibroblasts (HF57) were a generous gift from Steven Bates (Department of Biology, City of Hope).

Tissue Processing and Immunohistochemistry

After tissues were handled for pathologic-anatomical diagnosis, samples were fixed immediately in 85% ethanol supplemented with 0.2 mmol/L EDTA for 24 hours and embedded in paraffin thereafter. Six-micron serial sections of ethanol (EDTA)-fixed, paraffin-embedded tissues were mounted on Probe-On slides (Ventana Medical Systems, Tucson, Arizona) and baked at 56° C for 1 hour. Slides were deparaffinized by incubating the slides in xylene for 2 × 10 minutes and rehydrating in 100% ethanol for 2 × 5 minutes, in 96% ethanol for 2 × 5 minutes, and in 70% ethanol for 2 × 5 minutes. Antigen retrieval was performed by placing the slides into steam heat for 5 minutes. Antibodies used for double staining were PCNA (Ab-1) (Oncogene Research Products, Cambridge, Massachusetts) and NCL-p53-DO7 (Novocastra, Newcastle-upon-Tyne, United Kingdom). Staining was performed using a Biotech Techmate 1000 Immunostainer (Ventana Medical Systems). Double staining of slides was performed using a modified ABC-Method (Ventana Medical Systems). Human PCNA was stained brown using the chromogen 3'3-diaminobenzidine, and the p53 protein was stained red using the chromogen Bio-Red. Appropriate positive tissue controls and negative controls for primary antibodies were included in

every experiment. Tissue sections from a lung adenocarcinoma exhibiting p53 protein overexpression of tumor cells served as positive control. Staining without primary antibody was performed on the lung cancer tissue sections for demonstration of specificity of staining. Stained tissue sections were covered with aqueous mounting medium (Aquatex; Merck, Darmstadt, Germany) and a glass coverslip. Samples were stored frozen at −20° C until examined.

For determination of optimal tissue fixation and immunohistochemistry, fresh tissue samples were cut in three pieces and fixed in 95% ethanol/0.01 mmol/L EDTA, 85% ethanol/0.2 mmol/L EDTA, and 70% ethanol/1 mmol/L EDTA, respectively. After 24 hours, tissues were paraffin embedded, and immunohistochemical staining was performed as described above. The quality and reproducibility of staining was assessed independently by two pathologists (YS and LMW). DNA integrity was routinely assessed after Proteinase K digestion (2 mg/ml final concentration for 16 hours followed by a 10-minute inactivation step at 90° C) of cells from a variety of tissues scraped from serial sections taken at different stages of tissue processing and immunohistochemical staining. Eight-microliter aliquots of digested cell samples were electrophoresed through a 1% agarose gel for 30 minutes and stained with ethidium bromide.

Preparation of Single Cells from Cell Cultures and from Paraffin-Embedded, Immunohistochemically Stained Tissue Sections

Fibroblasts resuspended in PBS were spread onto glass slides and were air dried for 2 hours. After staining with methylene blue, single cells were picked with sterile needles (Microlance; Becton Dickinson, Franklin Lakes, New Jersey). Single cells or small cell clusters on tissue slides staining positive for p53 protein overabundance were microdissected using a manual micromanipulator (MP285; Sutter Instruments, Novato, California) equipped with a Tungsten needle (FHC Inc., Bowdoinham, Massachusetts). An inverted microscope (Nikon TMS, Melville, New York) was used. Microdissection was performed as follows: with a fine surgical needle and at lower magnification, the target cell was chosen and the area around it was freed from surrounding tissue by scraping a free space as close as possible to the target. After changing to a magnification of ×40, the cell was then pushed into the free space with the joystick-operated micromanipulator, picked up with a new surgical needle, and transferred into the PCR tube. Approximately 5 to 10 minutes are necessary to microdissect a single cell from normal colonic epithelium. Fibroblasts in a single-cell layer were dissected in less than 2 minutes per cell and could be dissected manually without the micromanipulator. Cells were transferred into 10 μl of 1× *Taq* PCR buffer No. 3 without $MgCL_2$ (Roche) containing 400 μg/ml (fibroblasts) or 1 mg/ml (single cells from tissue sections) of Proteinase K and 3% Tween-20 (Roche). Cell lysis was performed by incubation for 3 hours (fibroblasts) or 16 hours (single cells from tissue sections) at 50° C followed by a 10-minute inactivation at 90° C.

PCR

A recently described protocol for improved primer extension preamplification (I-PEP)-PCR (Dietmaier et al, 1999) was used with major modifications (Table 1) in a Perkin Elmer 9600 thermocycler (Foster City, California). I-PEP-PCR was set up by adding 30 μl of I-PEP mix [final concentration: 0.05 mg/ml gelatin, 20 μmol/L $(N)_{15}$ random primer, 1 mmol/L each dNTP, 3 U of *Taq* Expand High Fidelity polymerase, 6 mmol/L $MgCl_2$, in 1× PCR buffer No. 3 from Roche] to 10 μl of lysis buffer containing the lysed cell. PCR was run for 50 cycles as follows: Step 1: 95° C for 2 minutes; Step 2: 95° C for 30 seconds; Step 3: 28° C for 90 seconds; Step 4: ramp 0.1° C per second to 55° C; Step 5: 55° C for 2 minutes; Step 6: 68° C for 3 minutes; Step 7: go to Step 2, 49 times; Step 8: 68° C for 15 minutes; and Step 9: 4° C. Exons 5 to 9 of the *TP53* gene were amplified in a nested PCR approach: first-round PCR was performed as a multiplex PCR using upstream and downstream primers specific for each exon tagged with an unrelated chimeric 20-nucleotide sequence at the 5' end of all primers (Shuber et al, 1995) (Table 2). Second round PCR was performed in a nested or hemi-nested condition using primers given in Table 2. Five microliters of I-PEP-PCR product was aliquoted to 25 μl of the specific PCR first-round mix: 0.2 mmol/L each dNTP; 1.5 mmol/L $MgCl_2$; 0.4 μmol/L primers exons 5, 7, and 9; 0.6 μmol/L primers exons 6 and 8; 0.5 U of *Taq* Expand High Fidelity Polymerase; and 0.5 U of Anti-*Taq* Antibody (Gibco, Eggenstein, Germany). Each exon was amplified separately in a second-round PCR. Three microliters of first-round PCR product was aliquoted to 22 μl of mix containing 1.5 mmol/L $MgCl_2$; 0.2 mmol/L each dNTP; 0.4 mmol/L primers exons 5 and 7; 0.6 mmol/L primers exons 6, 8, and 9; and 0.3 U of *Taq* Expand High Fidelity Polymerase. Cycling conditions were identical for first-round and second-round PCR: Step 1: 95° C for 2 minutes; Step 2: 95° C for 30 seconds; Step 3: 55° C for 45 seconds; Step 4: 72° C for 1 minute; Step 5: go to step 2, 29 times; Step 6: 72° C for 10 minutes; and Step 7: 4° C. Negative control reactions with water instead of DNA were included in each experiment. The presence and relative quantity of PCR product was ascertained by resolution on a 2% agarose gel. The primers of the nested PCR reaction were used for DNA sequencing.

DNA Sequencing

PCR products were purified using Microcon-100 Microconcentrators (Amicon Inc., Beverly, Massachusetts) and sequenced using the Big Dye Terminator Sequencing Kit (Applied Biosystems, Foster City, California) according to the manufacturers instructions. Sequence changes were confirmed by reamplifying from the original I-PEP-PCR and sequencing in the opposite direction. Repeated amplifications of the exons from another PEP

aliquot of the same cell were performed to show that the mutation could be reproducibly identified in independent aliquots of the PEP. Sample chromatograms were analyzed in a blinded manner by two individuals. Seq Ed (PE Biosystems Foster City, California) and Sequencher software (Gene Codes, Inc., Ann Arbor, Michigan) were used to generate two or multiple sequence alignments, respectively. Differences between sequences were compared with the published wild-type sequence using the IARC *TP53* mutation database (http://www.iarc.fr/p53/).

Acknowledgements

We thank Dr. Wolfgang Dietmaier for helpful discussions, Mrs. Wenyan Li, Mrs. Martha Magallanes, Mrs. Xuemin Li and Mrs. Sofia Loera for expert technical assistance, and Dr. S. Bates for the cells.

References

Arnold MM, Srivastava S, Fredenburgh J, Stockard CR, Myers RB, and Grizzle WE (1996). Effects of fixation and tissue processing on immunohistochemical demonstration of specific antigens. Biotech Histochem 71:224–230.

Barnes DM, Hanby AM, Gillett CE, Mohammed S, Hodgson S, Bobrow LG, Leigh IM, Purkis T, MacGeoch C, and Spurr NK (1992). Abnormal expression of wild type p53 protein in normal cells of a cancer family patient. Lancet 340:259–263.

Bassarova AV and Popov AA (1998). Immunohistochemical detection of p53: Effect of fixation and methods of antigen retrieval. Folia Histochem Cytobiol 36:127–132.

Becker I, Becker KF, Rohrl MH, Minkus G, Schutze K, and Hofler H (1996). Single-cell mutation analysis of tumors from stained histologic slides. Lab Invest 75:801–807.

Dietmaier W, Hartmann A, Wallinger S, Heinmoller E, Kerner T, Endl E, Jauch KW, Hofstadter F, and Ruschoff J (1999). Multiple mutation analyses in single tumor cells with improved whole genome amplification. Am J Pathol 154:83–95.

Duddy SK, Gorospe S, and Bleavins MR (1998). Genetic analysis of multiple loci in microsamples of fixed paraffin-embedded tissue. Toxicol Sci 46:317–323.

El-Hashemite N and Delhanty JD (1997). A technique for eliminating allele specific amplification failure during DNA amplification of heterozygous cells for preimplantation diagnosis. Mol Hum Reprod 3:975–978.

Emmert-Buck MR, Bonner RF, Smith PD, Chuaqui RF, Zhuang Z, Goldstein SR, Weiss RA, and Liotta LA (1996). Laser capture microdissection [see comments]. Science 274: 998–1001.

Findlay I, Matthews P, and Quirke P (1998). Multiple genetic diagnoses from single cells using multiplex PCR: Reliability and allele dropout. Prenat Diagn 18:1413–1421.

Findlay I, Ray P, Quirke P, Rutherford A, and Lilford R (1995). Allelic drop-out and preferential amplification in single cells and human blastomeres: Implications for preimplantation diagnosis of sex and cystic fibrosis. Hum Reprod 10:1609–1618.

Forrester K, Lupold SE, Ott VL, Chay CH, Band V, Wang XW, and Harris CC (1995). Effects of p53 mutants on wild-type p53-mediated transactivation are cell type dependent. Oncogene 10:2103–2111.

Garvin AM, Holzgreve W, and Hahn S (1998). Highly accurate analysis of heterozygous loci by single cell PCR. Nucleic Acids Res 26:3468–3472.

Greenblatt MS, Bennett WP, Hollstein M, and Harris CC (1994). Mutations in the p53 tumor suppressor gene: Clues to cancer etiology and molecular pathogenesis. Cancer Res 54:4855–4878.

Hachiya M, Chumakov A, Miller CW, Akashi M, Said J, and Koeffler HP (1994). Mutant p53 proteins behave in a dominant, negative fashion in vivo. Anticancer Res 14:1853–1859.

Hahn S, Garvin AM, Di Naro E, and Holzgreve W (1998). Allele drop-out can occur in alleles differing by a single nucleotide and is not alleviated by preamplification or minor template increments. Genet Test 2:351–355.

Hahn S, Zhong XY, Troeger C, Burgemeister R, Gloning K, and Holzgreve W (2000). Current applications of single-cell PCR. Cell Mol Life Sci 57:96–105.

Huusko P, Castren K, Launonen V, Soini Y, Paakkonen K, Leisti J, Vahakangas K, and Winqvist R (1999). Germ-line TP53 mutations in Finnish cancer families exhibiting features of the Li-Fraumeni syndrome and negative for BRCA1 and BRCA2. Cancer Genet Cytogenet 112:9–14.

Karlsen F, Kalantari M, Chitemerere M, Johansson B, and Hagmar B (1994). Modifications of human and viral deoxyribonucleic acid by formaldehyde fixation. Lab Invest 71:604–611.

May P and May E (1999). Twenty years of p53 research: Structural and functional aspects of the p53 protein [published erratum appears in Oncogene 2000 Mar 23;19(13): 1734]. Oncogene 18:7621–7636.

Milner J and Medcalf EA (1991). Cotranslation of activated mutant p53 with wild type drives the wild-type p53 protein into the mutant conformation. Cell 65:765–774.

Persson AE, Ling G, Williams C, Backvall H, Ponten J, Ponten F, and Lundeberg J (2000). Analysis of p53 mutations in single cells obtained from histological tissue sections. Anal Biochem 287:25–31.

Ponten F, Williams C, Ling G, Ahmadian A, Nister M, Lundeberg J, Ponten J, and Uhlen M (1997). Genomic analysis of single cells from human basal cell cancer using laser-assisted capture microscopy. Mutat Res 382:45–55.

Ray PF and Handyside AH (1996). Increasing the denaturation temperature during the first cycles of amplification reduces allele dropout from single cells for preimplantation genetic diagnosis. Mol Hum Reprod 2:213–218.

Rechitsky S, Freidine M, Verlinsky Y, and Strom CM (1996). Allele dropout in sequential PCR and FISH analysis of single cells (cell recycling). J Assist Reprod Genet 13:115–124.

Roehrl MH, Becker KF, Becker I, and Hofler H (1997). Efficiency of single-cell polymerase chain reaction from stained histologic slides and integrity of DNA in archival tissue. Diagn Mol Pathol 6:292–297.

Schutze K and Lahr G (1998). Identification of expressed genes by laser-mediated manipulation of single cells. Nat Biotechnol 16:737–742.

Shuber AP, Grondin VJ, and Klinger KW (1995). A simplified procedure for developing multiplex PCRs. Genome Res 5:488–493.

Snabes MC, Chong SS, Subramanian SB, Kristjansson K, DiSepio D, and Hughes MR (1994). Preimplantation single-

cell analysis of multiple genetic loci by whole-genome amplification. Proc Natl Acad Sci USA 91:6181–6185.

Suarez-Quian CA, Goldstein SR, Pohida T, Smith PD, Peterson JI, Wellner E, Ghany M, and Bonner RF (1999). Laser capture microdissection of single cells from complex tissues. BioTechniques 26:328–335.

Varley JM, McGown G, Thorncroft M, James LA, Margison GP, Forster G, Evans DG, Harris M, Kelsey AM, and Birch JM (1999). Are there low-penetrance TP53 alleles? Evidence from childhood adrenocortical tumors. Am J Hum Genet 65:995–1006.

Villadsen R, Nielsen KV, Bolund L, and Briand P (2000). Complete loss of wild-type TP53 in a nontransformed human epithelial cell line is preceded by a phase during which a heterozygous TP53 mutant effectively outgrows the homozygous wild-type cells. Cancer Genet Cytogenet 116:28–34.

Wallace-Brodeur RR and Lowe SW (1999). Clinical implications of p53 mutations. Cell Mol Life Sci 55:64–75.

Walsh PS, Erlich HA, and Higuchi R (1992). Preferential PCR amplification of alleles: Mechanisms and solutions. PCR Methods Appl 1:241–250.

Wells D, Sherlock JK, Handyside AH, and Delhanty JD (1999). Detailed chromosomal and molecular genetic analysis of single cells by whole genome amplification and comparative genomic hybridisation. Nucleic Acids Res 27:1214–1218.

Williams C, Ponten F, Moberg C, Soderkvist P, Uhlen M, Ponten J, Sitbon G, and Lundeberg J (1999). A high frequency of sequence alterations is due to formalin fixation of archival specimens. Am J Pathol 155:1467–1471.

Yagi N, Satonaka K, Horio M, Shimogaki H, Tokuda Y, and Maeda S (1996). The role of DNase and EDTA on DNA degradation in formaldehyde fixed tissues. Biotech Histochem 71:123–129.

Zhang L, Cui X, Schmitt K, Hubert R, Navidi W, and Arnheim N (1992). Whole genome amplification from a single cell: Implications for genetic analysis. Proc Natl Acad Sci USA 89:5847–5851.

Environmental and Molecular Mutagenesis 39:216–227 (2002)

p53 as a Mutagen Test in Breast Cancer

Kathleen A. Hill and Steve S. Sommer*

Department of Molecular Genetics, City of Hope National Medical Center, Duarte, California

The *p53* gene is mutated in about half of all tumors. The *p53* gene can be used as a "mutagen test," that is, the relative frequencies of the different types of mutation can be used as an epidemiological tool to explore the contribution of exogenous mutagens vs. endogenous processes in particular cancers. *p53* has been used as a mutagen test in breast cancer. Surprisingly, the pattern of *p53* mutations differs among 15 geographically and ethnically diverse populations. In contrast, mutation patterns in the human *factor IX* gene are similar in geographically and ethnically diverse populations. Diverse *p53* mutation patterns in breast cancer are consistent with a significant contribution by a diversity of exogenous mutagens. Breast tissue may be uniquely sensitive to lipophilic mutagens because of its unique architecture, characterized by tiny islands of cancer-prone mammary epithelial cells surrounded by a sea of adipocytes. Mammary epithelial cells may be differentially susceptible to released lipophilic mutagens preferentially concentrated in adjacent adipocytes and originating in the diet. To test this hypothesis, we developed a method for measuring mutation load from ethanol-fixed, paraffin-embedded human tissues immunohistochemically stained with anti-*p53* antibodies. Single cells staining positively for *p53* overabundance are microdissected and the gene is sequenced. It is possible to identify individuals with a high mutation load in normal breast tissue and who are presumably at increased risk for breast cancer. In addition, analysis of the *p53* gene with appropriate mutation detection methodology markedly improves the prediction of early recurrence, treatment failure, and death in breast cancer patients. Mutagen tests and mutation load measurements are useful tools to identify the role of mutagens in breast cancer. Environ. Mol. Mutagen. 39: 216–227, 2002.
© 2002 Wiley-Liss, Inc.

Key words: ***p53* gene; environmental mutagens; endogenous processes; spontaneous mutations; mutagen test; mutation pattern; mutation load; mutation scanning; genetic predisposition**

BREAST CANCER INCIDENCE, EPIDEMIOLOGY, AND ETIOLOGY

Breast cancer is the most frequent malignancy occurring in women worldwide, with some 600,000 new cases arising each year, and accounts for nearly 20% of all cancers among women [Higginson et al., 1992]. Breast cancer is the major cause of cancer-related deaths in women residing in the Western world [Coleman, 2000; Kaplan and Wingard, 2000]. The worldwide incidence of breast cancer varies by at least fourfold [Coleman et al., 1993; Coleman, 2000]. The incidence and mortality from breast cancer varies significantly in ethnically and geographically distinct populations. Northern peoples, including Scandinavians, Scots, Welsh, English, and North Americans, have the highest rates of breast cancer, whereas Italians, Spaniards, and Central and South Americans have lower rates of breast cancer [Boring et al., 1994]. Striking evidence for the geographic-specific factors associated with breast cancer incidence come from altered breast cancer incidence for emigrants from low-risk populations to high-risk populations [Stemmermann, 1991; Coleman et al., 1993; Coleman, 2000]. Asians, in general, have had very low rates of breast cancer in the past, which increase upon immigration to the West [Ziegler et al., 1993]. This effect is seen in the first generation and is more dramatic in second and subsequent generations of Japanese in the United States. Despite intensive study, the origins of sporadic breast cancer are largely unknown [Friedenreich and Marrett, 2001]. Studies examining the role of specific exposures in breast cancer have found either inconsistent or weak associations [Millikan et al., 1995]. It would be predicted that epidemiologic studies would be inconsistent if different mutagens predominate in different high-risk co-

© 2002 Wiley-Liss, Inc.

Presented as a keynote address at the international symposium on Breast Cancer and Environmental Mutagens, Bridging Molecular Research to Medicine and Public Health, Durham, North Carolina, 24 September 2001.

Grant sponsor: National Institutes of Health; Grant numbers: NS33354 and HL39762.

*Correspondence to: Steve S. Sommer, Chair, Departments of Molecular Genetics and Molecular Diagnosis, Professor and Chair, Division of Human Genetics, City of Hope National Medical Center, 1500 E. Duarte Road, Duarte, CA 91010. E-mail: sommerlab@coh.org

Received 25 September 2001; provisionally accepted 7 November 2001; and in final form 23 November 2001

horts. Analysis of the pattern of mutation in an endogenous human gene generally should allow high-risk cohorts with different mutagen exposures to be distinguished from each other, given the different signatures of mutagens.

SIGNATURE MUTATIONS PERMIT IDENTIFICATION OF MUTAGENS

The mutational process can be characterized by the resulting pattern of mutation, defined as the relative frequencies of various categories of mutation (e.g., transitions, transversions, deletions, and insertions; see Table I). Virtually all studied chemical mutagens alter the mutation pattern in specific ways [Friedberg et al., 1995]. Spontaneous mutation patterns in transgenic and endogenous targets are distinct from mutation patterns following exposures to known mutagens. For example, benzo[a]pyrene (B[a]P) and ethylnitrosourea (ENU) specifically cause point mutations at G:C and A:T base pairs, respectively [Skopek et al., 1996; Walker et al., 1996]. Dimethylbenzanthracene (DMBA) causes primarily base pair substitutions, predominantly A:T → T:A (44%) and G:C → T:A (24%) [Manjanatha et al., 2000]. In human populations, signature mutations associated with specific exogenous mutagens have been observed in samples of somatic mutations in the *p53* tumor-suppressor gene [Harris and Hollstein, 1993; Hussain and Harris, 1999], for example, G → T transversions in lung cancers of cigarette smokers [Chiba et al., 1990], C → T and CC → TT tandem dipyrimidine transitions in skin cancers associated with excessive sun exposure [Brash et al., 1991], and G → T transversions at arginine 249 in liver cancers associated with aflatoxin exposure [Bressac et al., 1991; Hsu et al., 1991; Scorsone et al., 1992; Aguilar et al., 1993; Li et al., 1993]. Such signature mutations can be detected in the mutation pattern with high sensitivity. In a comparison of the pattern of *p53* mutations in the lung cancers of cigarette smokers with that in colon cancer and the germline, a significant difference ($P < 0.0001$) could be detected, even though only nine mutations in lung cancers were available [Sommer, 1990].

In the previous examples, one specific mutagen is believed to have had a dominant influence on the mutation pattern. In general, differences in the mutation patterns between populations are attributable to differences in both endogenous processes and exogenous mutagen exposures, and interactions between these influences. Furthermore, mutagen exposure may generally involve a mix of mutagens with comparable degrees of influence on the mutation pattern, rather than one dominant mutagen. Differences between the mutation patterns in population samples resulting from such complex differences can be detected.

p53 AS A "MUTAGEN TEST"

Alterations of the *p53* tumor-suppressor gene are the most frequent genetic abnormalities in human malignancies [Greenblatt et al., 1994]. The frequency of *p53* mutations in various cancer tissues ranges from 10 to 80% [Soussi et al., 1994]. The precise mechanisms by which the p53 protein carries out its biological function are not clear. The ability of p53 to bind to specific DNA sequences and to activate transcription suggests an important role in the regulation of cell proliferation [Harris, 1996a,b]. Loss of *p53* function eliminates the growth-arrest response to DNA damage [Kastan et al., 1992] and enhances the frequency of gene amplification [Livingstone et al., 1992; Yin et al., 1992], suggesting a role for *p53* in the control of a cell-cycle checkpoint and maintenance of the integrity of the genome [Lane, 1992]. *p53* controls the G1 cell cycle checkpoint by activating transcription of genes with *p53*-responsive elements. The products of these genes trigger G1 arrest after DNA damage, preventing replication of a damaged DNA template. The p53 protein is hypothesized to function in the maintenance of genomic stability by blocking cells from replication after DNA damage until the damage is repaired, or if the damage is extensive, by initiating apoptosis [Lane, 1992].

p53 has several key advantages as a model mutation reporter: (1) Mutations in *p53* are clustered within exons 5 to 8 (80% of mutations within exons 5 to 8; [Hartmann et al., 1995a]). (2) p53 protein is stabilized by missense mutations primarily occurring in exons 5 to 9. p53 protein stabilization permits immunohistochemical staining to identify cells with mutant *p53* [Bartek et al., 1990, 1991, 1993; Prosser et al., 1990; Davidoff et al., 1991; Varley et al., 1991]. (3) Mutation screening and DNA sequence determination of exons 5 to 9 involve reasonable effort [Kovach et al., 1991; Blaszyk et al., 1995]. (4) A great deal of information on the degree of sequence conservation is available [Soussi et al., 1990]. Despite speculations that the *p53* gene is "the guardian of the genome" [Lane, 1992], compelling data indicate that the frequency and pattern of mutation are generally unchanged in normal and tumor tissues of *p53* nullizygous mice that develop cancer early [Sands et al., 1995; Buettner et al., 1996, 1997]. This evidence suggests that *p53* is a good mutation reporter because the background rate of mutations in *p53*-deficient cells is similar to that of wild type cells.

ANALYSIS OF *p53* MUTATIONS IN BREAST CANCER

Analysis of mutations in the *p53* gene in breast carcinomas was developed as a model system for characterizing somatic mutations in very small samples of cellulary heterogeneous human cancer tissue [Kovach et al., 1991]. To avoid limitations to molecular analysis resulting from the admixture of normal and malignant cells in most human cancer, a method was developed to characterize mutations in genomic DNA from clusters of as few as 10 tumor cells in stained touch preparations of fresh or previously frozen tissue from primary and metastatic cancers. The touch preparations viewed by light microscopy contained only malignant cells in compact clusters. The morphology of cells stained with toluidine blue/methylene blue was adequate for

TABLE I. Mutations in Exons 5 to 9 of the *p53* Gene in Breast Cancer From 15 Populations

Origin of patients	No. of mutations	Percentage of tumors with mutations	Classes of mutation (%)							
			Deletion, insertion	Transitions at CpG G:C → A:T	Transitions not at CpG			Transversions		
					G:C → A:T	A:T → G:C	G:C → C:G	G:C → T:A	A:T → C:G	A:T → T:A
Low cancer risk										
Aomori, Japan[a]	15	56	26	20	20	7	13	7	7	0
Tokushima, Japan[b]	26	23	19	35	11	8	4	4	8	11
Sapporo, Japan[c]	17	71	0	41	17	18	6	0	12	6
Tokyo, Japan[d]	29	25	21	17	21	24	0	3	7	7
Kagoshima, Japan[e]	13	26	0	39	8	15	0	15	8	15
High cancer risk										
US midwest white[f]	29	30	38	21	21	3	7	0	3	7
Detroit black[g]	16	34	6	31	13	32	6	6	6	0
New Orleans black/white[h]	14	15	14	14	57	15	0	0	0	0
San Francisco, white[i]	14	24	7	14	7	28	28	21	0	0
San Diego, white[j]	31	31	13	13	19	23	19	3	7	3
Tennessee, white[k]	18	22	0	11	6	6	33	28	0	16
Scandinavia[l]	50	18	18	32	20	10	2	12	6	0
Scotland[m]	41	30	7	15	17	10	10	22	14	5
France[n]	30	19	10	27	20	0	10	16	7	10
Austria[o]	14	23	14	36	14	0	7	0	0	29

[a]Hartmann et al., 1996; [b]Blaszyk et al., 1996; Sasa et al., 1993; [c]Blaszyk et al., 1996; [d]Tsuda et al., 1993; [e]Umekita et al., 1994; [f]Saitoh et al., 1994; Kovach et al., 1996; [g]Blaszyk et al., 1994; [h]Shiao et al., 1995; [i]Deng et al., 1994; [j]Glebov et al., 1994; [k]Caleffi et al., 1994; [l]Anderson et al., 1993; Thorlacius et al., 1993; [m]Coles et al., 1992; [n]Mazars et al., 1992; Faille et al. 1994; [o]Hartmann et al., 1995.

TABLE II. Analysis of the Patterns of Described *p53* Mutations in Breast and Lung Cancer[a]

Cancer	No. of cohorts	No. of mutations	n × 8 table *P* value	No. of comparisons	$P < 0.05$ expected	$P < 0.05$ observed	$P < 0.10$ expected	$P < 0.10$ observed
Breast	15	357	0.0032	105	5.25	29	10.5	46
Lung	9	332	0.65	45	2.25	4	4.5	8

[a]Blaszyk et al., 1996.

the unequivocal determination of the malignant phenotype of all cells in clusters selected for study. A nested pair of PCR amplifications of DNA was used to increase the amount of target-gene sequence sufficiently to permit direct sequencing of regions of likely functional significance in a 1.9-kilobase genomic segment including exons 5 through 9 of *p53*. The exons encompass all the hot spots for *p53* mutations and cover the region in which the majority of mutations have been found [Hartmann et al., 1995a]. The touch preparation technique was performed for breast as a model system and was also successful in prostate, pancreas, ovary, colon, and thyroid [Kovach et al., 1991].

INCIDENCE OF *p53* MUTATIONS IN BREAST CANCER

A significant proportion of breast tumors contains cells with mutations in *p53* (Table I). Significant differences exist among human populations in the percentage of breast tumors with mutations in *p53*. The percentage of tumors with mutations in an examination of 15 populations including both populations with high and low breast cancer risk ranges from 15 to 71%. Both high- and low-risk populations have high and low incidences of *p53* mutations in breast tumors. The incidence of *p53* mutations (exons 5 to 8) in breast cancer is generally lower than the occurrence of mutations in the same region in colorectal tumors (86% of 28 tumors with loss of one *p53* allele) [Baker et al., 1990].

p53 MUTATION PATTERNS VARY AMONG 15 POPULATIONS

The majority of 15 breast cancer populations investigated show distinct patterns of mutation (Table I). Analyzing all 15 populations by the Fisher's exact test showed a global intergroup difference (Table II). The statistical significance is retained if any one of the 15 populations is excluded, indicating that no single sample accounted for the differences among populations. Pairwise analyses of the mutational patterns confirm an excess of differences among the populations (Table II). Many of the differences are substantial because statistical significance was often observed with only relatively few mutations divided among the eight mutational types. The differences are unlikely to be attributed to methodological biases because different patterns were found among six populations analyzed by one laboratory using identical methods [Blaszyk et al., 1994, 1996; Saitoh et al., 1994; Hartmann et al., 1995b, 1996]. The differences between populations with breast cancer are unlikely to represent varying mixtures of just two or three mutagens in diverse populations, given that subanalyses by the Fisher exact test reveal that the differences are attributed to highly distinct patterns of mutations (Table II).

Interestingly, low-risk populations from southern Japan showed an intermediate pattern of mutations, such that there was no significant difference from any of the high-risk populations [Blaszyk et al., 1996]. One hypothesis is that the mutation pattern in this population predominantly reflects a baseline endogenous pattern. Mutagens present in high-risk populations might skew the endogenous pattern in different directions, generating patterns that are more different from one another than from the endogenous mutation pattern. It will be of interest to analyze the pattern of *p53* gene mutations in other populations with a very low risk of breast cancer and in cohorts from low-risk populations who immigrated to a high-risk area.

DIFFERENCES IN *p53* MUTATION PATTERN AMONG U.S. POPULATIONS

Breast cancer in black American women has a worse prognosis than that in white women, especially in Midwest urban centers and the northeast in the United States. Black women in Michigan had the highest average annual age-adjusted mortality in the United States for 1984–1988. Global analyses of eight mutation types and three populations (black Detroit, rural Midwest white, and Scottish populations; Table III) show highly significant intergroup differences ($P = 0.014$, Fisher's exact test in a 3×8 table). Pairwise comparisons showed the major differences were a higher frequency of all types of transitions in U.S. blacks compared with that of U.S. whites ($P = 0.02$) and with Scottish whites ($P = 0.038$), and in particular, A:T → G:C transitions compared with U.S. whites ($P = 0.016$). Transitions at non-CpG dinucleotides were also more common in U.S. blacks than in whites ($P = 0.014$). G:C → T:A transversions were more common in the Scottish population than in U.S. whites ($P = 0.046$) and microdeletions/insertions were more frequent in the U.S. white group than in the U.S. blacks ($P = 0.042$) and in the Scots ($P = 0.006$).

The basis for the observed excess of A:T → G:C transitions (31%) in the black patients is not known but is hy-

TABLE III. Germline Factor IX Mutation Patterns Are Similar in Diverse Populations but Different From Somatic *p53* Mutation Patterns Associated With Mutagens

Selected mutation samples			Classes of mutation (%)							
				Transitions at CpG	Transitions not at CpG			Transversions		
Type	Origin of patients	No. of mutations	Deletion/ insertion	G:C → A:T	G:C → A:T	A:T → G:C	G:C → C:G	G:C → T:A	A:T → C:G	A:T → T:A
Germline *factor IX* mutations in population samples*	US/Canadian Caucasian[a]	127	16	28	15	17	6	10	6	4
	Mexican Hispanic[b]	32	19	34	22	6	6	3	6	3
	Chinese[c]	66	23	27	14	14	6	6	5	6
Somatic *p53* mutations in tumor samples	Lung cancer[d]	24	8	21	4	0	13	50	4	0
	Squamous cell skin cancer[e]	15	33	0	33	0	7	27	0	0
	Liver cancer[f]	13	8	0	0	0	8	85	0	0
Somatic *p53* mutations in breast tumor samples	Detroit blacks[g]	16	6	31	13	31	6	6	6	0
	Midwest US rural whites[h]	18	39	28	6	0	11	0	6	11
	Scottish whites[i]	41	7	15	17	10	10	22	15	5

*Swedish, German, French and Non-Chinese Asian (mostly Korean) and South African population samples also have similar mutation patterns (Montandon et al. 1992; Knobloch et al. 1993; Ghanem et al. 1993; Bottema et al. 1990; and Li et al. 2001, respectively)

Mutation data are taken from: [a]Gostout et al. 1993; [b]Thorland et al. 1995; [c]Liu et al. 2000; [d]Chiba et al. 1990; [e]Brash et al. 1992; [f]Bressac et al., 1991, Hsu et al. 1991; [g]Blaszyk et al. 1994; [h]Saitoh et al. 1994, Sommer et al. 1992; [i]Coles et al. 1992.

pothesized to be a population-specific environmental exposure or endogenous process. A:T → G:C transitions comprise 10% of all mutations reported in the *p53* gene in human cancer, including breast cancers from pooled populations [Biggs et al., 1993]. An excess of A:T → G:C mutations in a particular population might result from increased exposure to exogenous or endogenous nitric oxide and/or defective repair of the deaminated adduct induced by nitric oxide [Wink et al., 1991; Habuchi et al., 1993]. The detection of a novel fingerprint of mutations in breast cancers of black women leads to a testable hypothesis that an important bioregulatory molecule, nitric oxide, and other types of nitrites, might be significant contributors to mutation of the *p53* gene in certain populations.

p53 MUTATION PATTERNS IN BREAST, LUNG, SKIN, AND LIVER CANCER

In contrast to breast cancer, the number of pairwise differences found in lung cancer among nine cohorts of predominantly smokers is similar to chance expectation (Table II). Diverse ethnic groups might respond similarly to mutagens in cigarette smoke, or this powerful mutagen mixture overrides genetic predispositions in diverse populations. The pattern in smokers differs markedly from the pattern observed in lung cancers from nonsmokers [Takeshima et al., 1993]. Ethnically and/or geographically diverse populations show similar mutational responses to the complex mixture of mutagens present in cigarette smoke. *p53* mutation patterns in breast cancer are also significantly different from the *p53* mutation patterns in skin and liver cancer (Table IV).

THE HUMAN *FACTOR IX* GENE AS A GERMLINE "MUTAGEN TEST"

The observations in somatic mutation data indicate that any mutagen or mix of mutagens that has a substantial influence on germline mutation in a particular population would be revealed in the germline mutation pattern. Hence, the germline mutation pattern in the human *factor IX* gene has been analyzed in mutation samples from geographically, racially, and ethnically diverse human populations [Bottema et al., 1993; Saad et al., 1994; Sommer, 1994; Thorland et al., 1995; Chen et al., 1998; Heit et al., 1998; Morley and Turner, 1999; Li et al., 2000; Liu et al., 2000]. Because there was no selection for any particular mutagen exposure, these samples represent typical mutagen exposures in the population. The germline mutation patterns in these diverse population samples are significantly different from somatic *p53* mutation patterns discussed above but are remarkably similar to each other (Tables III and IV). The similarity of the observed germline mutation patterns suggests that if there is any population-specific mutagen exposure, it does not have a substantial influence on germline

TABLE IV. Pairwise Comparison of Mutation Patterns in the Factor IX and *p53* Genes[a]

		Germline factor IX mutation in population samples			Somatic *p53* mutations in tumor samples			Somatic *p53* mutations in breast tumor samples		
		US/Canadian Caucasian	Mexican Hispanic	Chinese	Lung cancer	Squamous cell skin cancer	Liver cancer	Detroit blacks	Midwest US rural whites	Scottish whites
Germline factor IX mutation in population samples	US/Canadian Caucasian		0.9	0.6	**0.001**	**0.004**	**0.001**	0.880	**0.043**	0.081
	Mexican Hispanic			0.6	**0.003**	**0.010**	**0.001**	0.402	0.430	**0.100**
	Chinese				**0.001**	**0.020**	**0.001**	0.598	0.426	**0.037**
Somatic *p53* mutations in tumor samples	Lung cancer					**0.004**	**0.040**	**0.004**	**0.001**	0.145
	Squamous cell skin cancer						**0.000**	**0.003**	**0.007**	0.078
	Liver cancer							**0.000**	**0.000**	**0.011**
Somatic *p53* mutations in breast tumor samples	Detroit blacks								**0.027**	0.361
	Midwest US rural whites									**0.012**
	Scottish whites									

[a]Pairwise comparison of the patterns in Table III. *P* values were computed using Fisher's exact test.

mutation. Detailed analysis of the human *factor IX* gene as a model for germline mutation reveals a pattern of mutation that is similar among a variety of ethnic and geographic groups [reviewed in Sommer et al., 2001]. The pattern of germline mutation in the *factor IX* gene indicates that ethnically and/or geographically diverse populations respond similarly to the heterogeneous mutational insults resulting from multiple endogenous germline mutagens.

The similarity of the observed germline mutation patterns is compatible with the hypothesis that universal endogenous mechanisms are the dominant causes of human germline mutation [Gostout et al., 1993]. This hypothesis makes sense from an evolutionary perspective because endogenous processes that cause germline mutation provide a selective advantage by enhancing the ability to adapt to environmental changes. It is likely that evolution selected endogenous processes that cause germline mutation to set the rate of germline mutation sufficiently high for adaptation to environmental changes, but not so high that it causes excessive morbidity. Endogenous mechanisms that may be involved in germline mutation include physicochemical processes, the action of endogenous molecules, and enzymatic processes [Bottema et al., 1991]. One physicochemical process, spontaneous deamination of methylated cytosines to thymines, is the likely cause of mutations that occur at a CpG dinucleotide, approximately one quarter of all *factor IX* mutations [Gojobori et al., 1982]. Certain mutagens, however, may preferentially affect CpG residues [Li et al., 1984].

Further support of the predominance of endogenous mechanisms in germline mutation comes from two lines of evidence that indicate that human germline mutation has not been substantially influenced by any recent increased exposures to mutagens. Increases in mutagen exposures in recent times might be expected as a consequence of the vast increases in the number and volume of man-made chemicals produced in modern industrial societies. The first line of evidence is that the pattern of mutations ascertained in hemophilia B patients and which originated in the last two to four generations is compatible with the pattern of mutation required to produce the bias against G and C nucleotides in *factor IX* (40% G + C content) that was shaped long before modern industrial times [Bulmer, 1986]. The second line of evidence is that the pattern of *factor IX* deep intronic sequence changes with respect to the *factor IX* sequence that existed when humans diverged from the other primates is similar to the pattern of recent mutations in hemophilia B patients [Feng et al., 2002]. These deep intronic sequence changes provide a window into the mutational processes before modern times, given that most occurred earlier and persisted because of the lack selective pressure in these regions. In the context of the hypothesis that endogenous mechanisms predominate in and essentially control germline mutation, these lines of evidence indicate that any increased exposure to mutagens associated with modern industrial societies has not overwhelmed the cellular control of germline mutation.

A POPULATION-SPECIFIC HOTSPOT OF MUTATION IN THE *FACTOR IX* GENE

Nucleotide 17747 was found to be a significant non-CpG hotspot of *factor IX* germline mutation in Latin American populations [Drost et al., 2000]. Observed recurrent sites of mutation in a population sample may be indicative of mutagenic mechanisms specific to that population, although the possibility that such an observation occurred by chance must be ruled out. Initial population samples may be small, so statistical methods appropriate for small samples must be applied. For example, two base substitutions at each of two nucleotides in *factor IX* but not part of CpG dinucleotides were recently reported in a small population sample collected in Mexico [Ketterling et al., 1994]. When these new data were combined with previously collected mutation data to form two progressively larger and inclusive Latin American samples, additional mutations were observed at one recurrent site, nucleotide 17747, and an additional recurrent nucleotide was observed. The recurrent nucleotides in these

samples were significant ($P < 0.0003$ for both combined sets as computed by simulation) [Sommer and Ketterling, 1993]. In contrast, in three non-Latin American control samples, there was at most only one nucleotide that recurred only once, most likely a chance recurrence ($P \geq 0.5$). When the significance of substitutions was individually analyzed at each recurrent nucleotide, nucleotide 17747 was shown to be a significant recurrent nucleotide by itself in all the Latin American population samples ($P \leq 0.02$). Furthermore, a standard statistical comparison of mutation frequencies in the previously collected data alone confirmed that the frequency of mutation at nucleotide 17747 is significantly higher in Latin Americans than in all other populations combined ($P = 0.01$). Thus, nucleotide 17747 is a germline mutation hotspot in *factor IX* specific to Latin American populations. This may be the first evidence for population-specific effects on germline mutation that causes human genetic disease. No dramatic population-specific hotspots of mutation in the *p53* gene in breast cancer were observed.

A ROLE FOR DIETARY FAT IN THE EPIDEMIOLOGY OF BREAST CANCER

It is unclear what role the amount and nature of dietary fat may play in the initiation and promotion of breast cancer. The results of studies of dietary fat and calories in experimental rodent mammary gland tumorigenesis show that hyperalimentation of fat, either saturated or unsaturated, significantly stimulates tumorigenesis [Welsch, 1994]. Mice fed a Western-style diet (high fat and phosphate and low calcium and vitamin D) show significantly increased mammary epithelial cell proliferation [Lok et al., 1992; Xue et al., 1996]. The stimulatory effect of high levels of dietary fat appears to act primarily at the promotional stage. In addition, fatty acids may affect initiation and promotion of breast cancer, perhaps related to the formation of reactive metabolites of unsaturated fatty acids (malondialdehyde) [Hietanen et al., 1986]. A current view is that the mammary tumorigenic-enhancing activities of a high-fat diet is, at least in part, through a caloric mechanism [Welsch, 1994]. Epidemiologic investigations regarding the relationship between the amount and the type of fat consumed and breast cancer have produced conflicting results [Hunter et al., 1996; Clavel-Chapelon et al., 1997; Wynder et al., 1997; Woutersen et al., 1999] and a mechanism by which dietary fat may modulate carcinogenesis has not been elucidated.

A likely interpretation of the mutation pattern data is that different mutagens produce breast cancer in different populations. Although direct evidence is lacking, there is growing indirect evidence that genotoxic agents in the environment and particularly of dietary origin may play a causative role in the etiology of breast cancer [Biggs et al., 1993; Ziegler et al., 1993]. Some of the agents are even known to be mammary carcinogens in rodents [El-Bayoumy, 1992]. Hydrophobic mutagens have been found to adsorb to cereal brans and cereal bran dietary fibers [Harris et al., 1998]. The anatomy of breast tissue may predispose mammary epithelial cells to mutagenesis by lipophilic mutagens present in adipocytes (70–90% breast mass). It has been postulated that lipid-soluble agents could be sequestered by mammary tissue [Beer and Billingham, 1978]. Throughout adult life, the glandular tissue of the breast in nonlactating women secretes and reabsorbs fluid [Petrakis et al., 1980]. Such events may result in exposure of epithelial cells to xenobiotics [El-Bayoumy, 1992].

Compounds of endogenous and/or exogenous origin have been detected in breast fluid and some of these compounds may be genotoxic [Wrensch et al., 1989]. Perhaps this reflects the unique biology of breast tissue, in which clusters of mammary cells are surrounded by adipocytes. In adulthood, adipocytes cannot replicate but they can concentrate the great diversity of lipophilic mutagens from vegetables and from meat products (which may well derive from the diets fed to the animals). These stored lipophilic mutagens could diffuse into neighboring mammary cells. This hypothesis is compatible with the frequent but inconsistent correlation of fat in the diet with risk of breast cancer [Byers, 1994]. Mammary epithelial cells may be: (1) differentially susceptible to lipophilic mutagens originating in the diet and preferentially concentrated in adjacent adipocytes and (2) differentially susceptible to the type of fat consumed in the diet. Forty percent of lipid extracts from human breast tissues (derived from reduction mammoplasties) are mutagenic in in vitro mutation assays [Martin et al., 1997, 1998a,b]. These lipid extracts also produced evidence of DNA damage in cultured cells isolated from human breast milk [Martin et al., 1999]. The anatomy of breast tissue may predispose human mammary epithelium to mutations from lipophilic mutagens concentrated in adjacent adipocytes. A sea of adipocytes surrounds tiny islands of cancer-prone mammary cells. Adipocytes may also be heavily mutagenized by lipophilic mutagens, but they are unable to divide, so cancer is not an outcome of these mutations.

MUTATION LOAD MEASURED IN NORMAL HUMAN TISSUES

To test the lipophilic mutagen hypothesis another approach is to determine mutation load (the frequency and pattern of mutations) in normal and tumor tissues in different populations exposed to different diets. Recent advances permit routine mutation analysis using a single cell microdissected from routinely processed paraffin-embedded normal and tumor tissues. Major improvements have recently been made to tissue fixation and processing and single cell DNA amplification to enable the analysis of rare heterozygous mutations in the *p53* gene of single cells microdissected from ethanol-fixed, paraffin-embedded, and immunohistochemically stained sections of normal human tissue [Heinmoller, 2002]. The tissue-fixation protocol for paraffin-embedded tissues has vastly improved the mainte-

TABLE V. Univariate Analysis of Prognostic Factors for Overall and Disease-Free Survival in 90 Breast Cancer Patients[a]

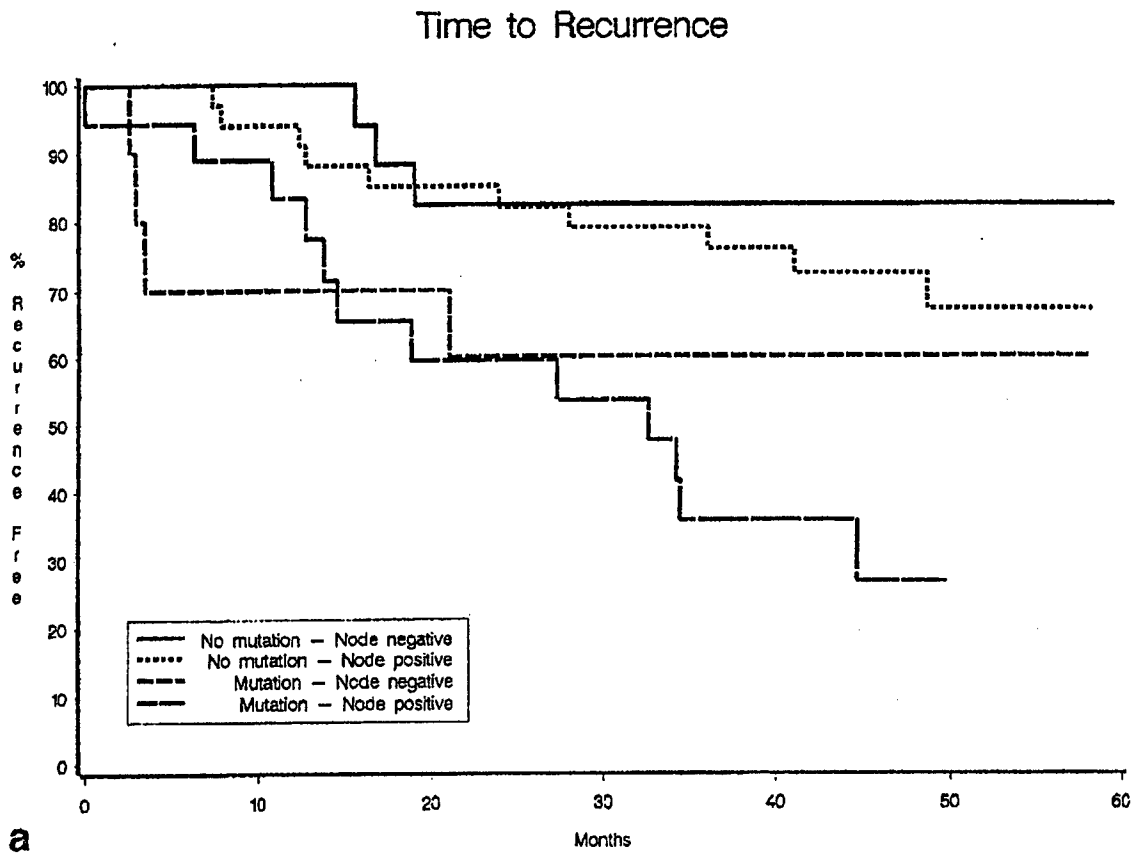
Prognostic Indicator	Disease-free survival (*P* value)[b]	Overall survival (*P* value)[b]
Estrogen receptor status (negative vs. positive)	0.002	0.053
Lymph node status (positive vs. negative)	0.177	0.209
Lymph node status (>4 pos. vs <4 pos.)	0.276	0.138
TNM tumor stage[c] (IIB–III vs. I + IIA)	0.095	0.703
Tumor size (>2 cm vs. <2 cm)	0.430	0.18
p53 IHC[d]	0.020	0.003
p53 mutation	0.003	0.0001

[a]Blaszyk et al., 2000. Average follow-up of 60 months.
[b]Exact log-rank statistics.
[c]American Joint Committee on Cancer (3rd ed.).
[d]IHC; immunohistochemistry.

nance of DNA integrity with the use of ethanol and EDTA. Single-cell microdissection was achieved by a simple and easy-to-operate manual microdissection device. The success rate of PCR amplification from DNA obtained from a single cell dissected from paraffin-embedded tissue has been increased to a level where serial analyses of mutations in a single cell from normal or tumor tissue can be performed in a reasonable number of experiments (Heinmoller et al. 2002; Sommer, unpublished observations). This technique has demonstrated preferential and significantly increased detection of mutations of likely functional relevance in single cells staining immunohistochemically for *p53* overabundance (Sommer, unpublished observations). This methodology is to be the basis for further studies to determine a *p53* mutation load in normal tissues in the context of cell and tissue morphology, enabling an estimation of individual cancer risk well in advance of the development of malignant disease.

p53 GENE MUTATION MAY BE THE MOST IMPORTANT PROGNOSTIC INDICATOR IN BREAST CANCER

The driving force for the work on *p53* as a mutagen test in human breast cancer was the use of mutations in this gene as an epidemiological tool to compare cohorts with similar mutagenic etiology, given the initial hypothesis that there might be two or three important etiologies in breast cancer rather than just one. The work as described suggests a diversity of mutagens as judged by the multiple patterns of mutation found in geographically or ethnically diverse populations. However, follow-up of the analyzed patients also allows the effect of *p53* mutations on outcome to be assessed. Cells lacking normal *p53* function have a selective growth advantage and are more resistant to ionizing radiation and some widely used anticancer drugs than cells with wild type p53 protein [Lowe et al., 1993; Zambetti and Levine, 1993]. Thus cancers with mutated *p53* genes might be expected to behave more aggressively clinically than tumors with preserved normal *p53* function. Initial studies of the *p53* gene in breast cancer examined the association between cancer prognosis and *p53* overexpression detected by immunohistochemistry [reviewed by Barbareschi, 1993], although differences in technique and variability in the frequency and intensity of the immunoreactivity confound these analyses. In a meta-analysis of *p53* overexpression using immunohistochemistry and over 9000 breast cancer patient tumors, the prognostic and predictive value of the p53 protein overexpression appeared weak [Barbareschi, 1993]. The prognostic significance of *p53* mutations in breast cancer was reported in an analysis of mutations detected by a yeast functional assay [Chappuis et al., 1999]. However, the yeast functional assay detects only a subclass of *p53* mutations and the functional relevance of these mutations in human cancers is not known. Examination of a subclass of *p53* mutations is expected to underestimate the significance of *p53* mutations as a prognostic indicator because false negatives are generally the major problem with a functional assay. Several other studies have directly correlated *p53* gene mutations with breast cancer prognosis but none was conducted prospectively [reviewed in Blaszyk et al., 2000]. Relative risks were between 2.2 and 3.3 but these analyses are difficult to compare because of different variables used for analyses by the different groups. These studies underestimate the contribution of *p53* mutations because significant numbers of mutations are missed [Hartmann, 1997]. Analyses were designed to detect virtually all mutations by use of the following methods: (1) touch preparations of breast cancers to yield clusters of pure tumor cells; (2) dideoxyfingerprinting [Sarkar et al., 1992], which detects virtually 100% of mutations in the *p53* gene and other genes [Sarkar et al., 1992; Blaszyk et al., 1995], to screen for mutations; and (3) analysis of exons 4 through 10 of the *p53* gene. *p53* mutations were found to be the single most important predictor for early recurrence and death (Table V) [Kovach et al., 1996; Blaszyk et al., 2000]. Substantial prognostic significance of *p53* mutations in



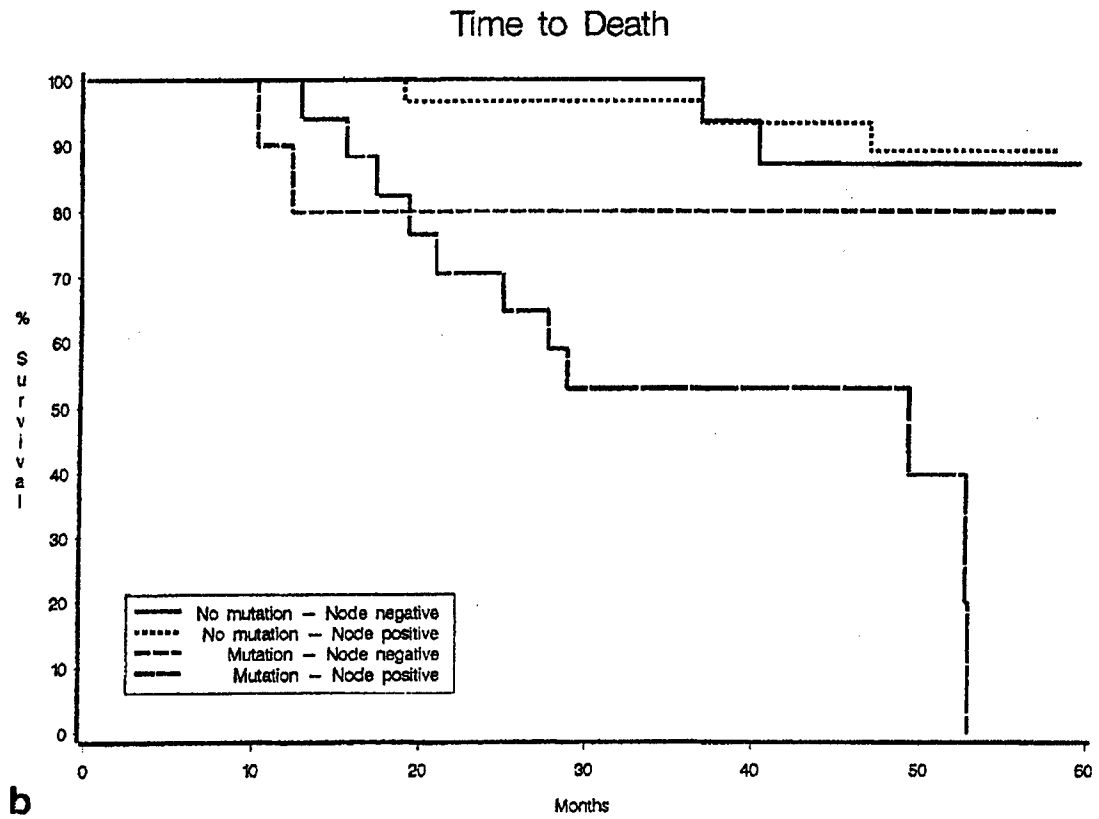


Fig. 1. Distribution of time to (a) breast cancer recurrence and (b) death in patients with positive (n = 52) and with negative (n = 27) axillary lymph nodes according to the presence vs. the absence of a *p53* gene mutation as estimated by the Kaplan–Meier method [Balszyk et al., 2000].

breast cancer remained significant with a follow-up period of 60 months (Fig. 1a, b) [Blaszyk et al., 2000].

CONCLUSION

In conclusion our present data demonstrate that *p53* mutation patterns are heterogeneous, consistent with population-specific environmental mutagen exposures. Two approaches are currently in use to test a hypothesis that mammary epithelial cells may be differentially susceptible to a diversity of lipophilic mutagens originating in the diet and preferentially concentrated in adjacent adipocytes. Spontaneous mutation patterns in mouse, including mutation patterns for mammary tissues, do not show tissue specificity and provide a baseline for analysis of environmental mutagen exposures. It is also possible to measure mutation load in normal human tissues. Use of methodology designed to detect virtually all mutations revealed that *p53* gene mutations might be the most important prognostic indicator in breast cancer.

REFERENCES

Aguilar F, Hussain SP, Cerutti P. 1993. Aflatoxin B_1 induces the transversion of G→T in codon 249 of the p53 tumor suppressor gene in human hepatocytes. Proc Natl Acad Sci USA 90:8586–8590.

Baker SJ, Preisinger AC, Jessup JM, Paraskeva C, Markowitz S, Willson JK, Hamilton S, Vogelstein B. 1990. p53 gene mutations occur in combination with 17p allelic deletions as late events in colorectal tumorigenesis. Cancer Res 50:7717–7722.

Barbareschi M. 1993. Prognostic value of the immunohistochemical expression of p53 in breast carcinomas: a review of the literature involving over 9,000 patients. Appl Immunohistochem 4:106–116.

Bartek J, Iggo R, Gannon J, Lane DP. 1990. Genetic and immunochemical analysis of mutant p53 in human breast cancer cell lines. Oncogene 5:893–899.

Bartek J, Bartkova J, Vojtesek B, Staskova Z, Lukas J, Rejthar A, Kovarik J, Midgley CA, Gannon JV, Lane DP. 1991. Aberrant expression of the p53 oncoprotein is a common feature of a wide spectrum of human malignancies. Oncogene 6:1699–1703.

Bartek J, Bartkova J, Lukas J, Staskova A, Vojtesek B, Lane DP. 1993. Immunohistochemical analysis of the p53 oncoprotein on paraffin sections using a series of novel monoclonal antibodies. J Pathol 169:27–31.

Beer AE, Billingham RE. 1978. Adipose tissue, a neglected factor in aetiology of breast cancer? Lancet 2:296.

Biggs PJ, Warren W, Venitt S, Stratton MR. 1993. Does a genotoxic carcinogen contribute to human breast cancer? The value of mutational spectra in unravelling the aetiology of cancer. Mutagenesis 8:275–283.

Blaszyk H, Vaughn CB, Hartmann A, McGovern RM, Schroeder JJ, Cunningham J, Schaid D, Sommer SS, Kovach JS. 1994. Novel pattern of p53 gene mutations in an American black cohort with high mortality from breast cancer. Lancet 343:1195–1197.

Blaszyk H, Hartmann A, Schroeder JJ, McGovern RM, Sommer SS, Kovach JS. 1995. Rapid and efficient screening for p53 gene mutations by dideoxyfingerprinting (ddF). BioTechniques 18:256–260.

Blaszyk H, Hartmann A, Tamura Y, Saitoh S, Cunningham JM, McGovern RM, Schroeder JJ, Schaid DJ. 1996. Molecular epidemiology of breast cancers in northern and southern Japan: the frequency, clustering, and patterns of p53 gene mutations differ among these two low-risk populations. Oncogene 13:2159–2166.

Blaszyk H, Hartmann A, Cunningham JM, Schaid D, Wold LE, Kovach JS, Sommer SS. 2000. A prospective trial of midwest breast cancer patients: a p53 gene mutation is the most important predictor of adverse outcome. Int J Cancer 89:32–38.

Boring CC, Squires TS, Tong T, Montgomery S. 1994. Cancer statistics, 1994. CA Cancer J Clin 44:7–26.

Bottema CDK, Bottema MJ, Ketterling RP, Yoon H-S, Janco RL, Phillips JA III, Sommer SS. 1991. Why does the human factor IX gene have a G + C content of 40%? Am J Hum Genet 49:839–850.

Bottema CDK, Ketterling RP, Vielhaber E, Yoon H-S, Gostout B, Jacobson DP, Shapiro A, Sommer SS. 1993. The pattern of spontaneous germline mutation: relative rates of mutation at or near CpG dinucleotides in the factor IX gene. Hum Genet 91:496–503.

Brash DE, Rudolph JA, Simon JA, Lin A, McKenna GJ, Baden HP, Halperin AJ, Ponten J. 1991. A role for sunlight in skin cancer: UV-induced p53 mutations in squamous cell carcinoma. Proc Natl Acad Sci USA 88:10124–10128.

Bressac B, Kew M, Wands J, Ozturk M. 1991. Selective G to T mutations of p53 gene in hepatocellular carcinoma from southern Africa. Nature 350:429–431.

Buettner VL, Hill KA, Nishino H, Schaid DJ, Frisk CS, Sommer SS. 1996. Increased mutation frequency and altered spectrum in one of four thymic lymphomas derived from tumor prone p53/Big Blue transgenic mice. Oncogene 13:2407–2413.

Buettner VL, Nishino H, Haavik J, Knöll A, Hill KA, Sommer SS. 1997. Spontaneous mutation frequencies and spectra in p53 (+/+) and p53 (−/−) mice: a test of the "guardian of the genome" hypothesis in the Big Blue transgenic mouse mutation detection system. Mutat Res 379:13–20.

Bulmer M. 1986. Neighboring base effects on substitution rates in pseudogenes. Mol Biol Evol 3:322–329.

Byers T. 1994. Nutritional risk factors for breast cancer [review]. Cancer 74:288–295.

Chappuis PO, Estreicher A, Dieterich B, Bonnefoi H, Otter M, Sappino AP, Iggo R. 1999. Prognostic significance of p53 mutation in breast cancer: frequent detection of non-missense mutations by yeast functional assay. Int J Cancer 84:587–593.

Chen JX, Zheng Y, West M, Tang MS. 1998. Carcinogens preferentially bind at methylated CpG in the p53 mutational hot spots. Cancer Res 58:2070–2075.

Chiba I, Takahashi T, Nau MM, D'Amico D, Curiel DT, Mitsudomi T, Buchhagen DL, Carbone D, Piantadosi S, Koga H. 1990. Mutations in the p53 gene are frequent in primary, resected non-small cell lung cancer. Lung Cancer Study Group. Oncogene 5:1603–1610.

Clavel-Chapelon F, Niravong M, Joseph RR. 1997. Diet and breast cancer: review of the epidemiologic literature. Cancer Detect Prev 21:426–440.

Coleman MP. 2000. Trends in breast cancer incidence, survival, and mortality [letter]. Lancet 356:590–591.

Coleman MP, Esteve J, Damiecki P, Arslan A, Renard H. 1993. Trends in cancer incidence and mortality. IARC Sci Publ 121:1–806.

Davidoff AM, Herndon JEII, Glover NS, Kerns B-JM, Pence JC, Iglehart JD, Marks JR. 1991. Relation between p53 overexpression and established prognostic factors in breast cancer. Surgery 110:259–264.

Drost JB, Scaringe WA, Jaloma-Cruz AR, Li X, Ossa DF, Kasper CK, Sommer SS. 2000. Novel hotspot detector software reveals a non-CpG hotspot of germline mutation in the factor IX gene (F9) in Latin Americans. Hum Mutat 16:203–210.

El-Bayoumy K. 1992. Environmental carcinogens that may be involved in human breast cancer etiology. Chem Res Toxicol 5:585–590.

Feng J, Drost JB, Scaringe WA, Liu Q, Sommer SS. 2002. The relative rates of ancient germline mutations during human evolution are

similar to those of recent germline mutations. Hum Mutat 19:49–57.

Friedberg EC, Walker GC, Siede W. 1995. DNA repair and mutagenesis. Washington, DC: ASM Press. 694 p.

Friedenreich C, Marrett LD. 2001. Workshop report: identification of research needs breast cancer etiology. Chronic Dis Can 22:41–49.

Gojobori T, Li WH, Graur D. 1982. Patterns of nucleotide substitution in pseudogenes and functional genes. J Mol Evol 18:360–369.

Gostout B, Vielhaber E, Ketterling RP, Yoon H-S, Bottema CDK, Kasper CK, Koerper M, Sommer SS. 1993. Germline mutations in the factor IX gene: a comparison of the pattern in Caucasians and non-Caucasians. Hum Mol Genet 2:293–298.

Greenblatt MS, Bennett WP, Hollstein M, Harris CC. 1994. Mutations in the p53 tumor suppressor gene: clues to cancer etiology and molecular pathogenesis. Cancer Res 54:4855–4878.

Habuchi T, Takahashi R, Yamada H, Ogawa O, Kakehi Y, Ogura K, Hamazaki S, Toguchida J, Ishizaki K, Fujita J, Sugiyama T, Yoshida O. 1993. Influence of cigarette smoking and schistosomiasis on p53 gene mutation in urothelial cancer. Cancer Res 53:3795–3799.

Harris CC. 1996a. Structure and function of the p53 tumor suppressor gene: clues for rational cancer therapeutic strategies. J Natl Cancer Inst 88:1442–1455.

Harris CC. 1996b. 53 tumor suppressor gene: from the basic research laboratory to the clinic—an abridged historical perspective. Carcinogenesis 17:1187–1198.

Harris CC, Hollstein M. 1993. Clinical implications of the p53 tumor-suppressor gene. N Engl J Med 329:1318–1327.

Harris PJ, Vallappilakkandy KS, Roberton AM, Triggs CM, Blakeney AB, Ferguson LR. 1998. Adsorption of a hydrophobic mutagen to cereal brans and cereal bran dietary fibres. Mutat Res 412:323–331.

Hartmann A, Blaszyk H, McGovern RM, Schroeder JJ, Cunningham J, DeVries EMG, Kovach JS, Sommer SS. 1995a. p53 gene mutations inside and outside of exons 5–8: the patterns differ in breast and other cancers. Oncogene 10:681–688.

Hartmann A, Rosanelli G, Blaszyk H, Cunningham JM, McGovern RM, Schroeder JJ, Schaid DJ, Kovach JS, Sommer SS. 1995b. Novel pattern of p53 mutation in breast cancers from Austrian women. J Clin Invest 95:686–689.

Hartmann A, Blaszyk H, Saitoh S, Tsushima K, Tamura Y, Cunningham JM, McGovern RM, Schroeder JJ, Sommer SS, Kovach JS. 1996. High frequency of p53 mutations in primary breast cancers in Japanese women, a low-incidence population. Br J Cancer 73:896–901.

Heinmoller E, Liu, Q, Sun, S, Schlake G, Hill, KA, Weiss LM, Sommer SS. 2002. Toward efficient analysis of mutations in single cells from ethanol-fixed, paraffin-emmbedded and immunohistochemically-stained tissues. Laboratory Investigations (in press).

Heit JA, Thorland EC, Ketterling RP, Lind TJ, Daniels TM, Zapata RE, Ordonez SM, Kasper CK, Sommer SS. 1998. Germline mutations in Peruvian patients with hemophilia B: pattern of mutation in AmerIndians is similar to the putative endogenous germline pattern. Hum Mutat 11:372–376.

Hietanen E, Punnonen K, Punnonen R, Auvinen O. 1986. Fatty acid composition of phospholipids and neutral lipids and lipid peroxidation in human breast cancer and lipoma tissue. Carcinogenesis 7:1965–1969.

Higginson J, Muir CS, Munoz N. 1992. Human cancer: epidemiology and environmental causes. In: Anonymous. Cambridge Monographs on Cancer Research. Cambridge: Cambridge Univ. Press. p 377–387.

Hsu IC, Metcalf RA, Sun T, Welsh JA, Wang NJ, Harris CC. 1991. Mutational hotspot in the p53 gene in human hepatocellular carcinomas. Nature 350:427–428.

Hunter DJ, Spiegelman D, Adami HO, Beeson L, van den Brandt PA, Folsom AR, Fraser GE, Goldbohm RA, Graham S, Howe GR. 1996. Cohort studies of fat intake and the risk of breast cancer—a pooled analysis [see comments]. N Engl J Med 334:356–361.

Hussain SP, Harris CC. 1999. p53 mutation spectrum and load: the generation of hypotheses linking the exposure of endogenous or exogenous carcinogens to human cancer. Mutat Res 428:23–32.

Kaplan RM, Wingard DL. 2000. Trends in breast cancer incidence, survival, and mortality [letter]. Lancet 356:592–593.

Kastan MB, Zhan Q, El-Deiry WS, Carrier F, Jacks T, Walsh WV, Plunkett BS, Vogelstein B, Fornace AJ Jr. 1992. A mammalian cell cycle checkpoint pathway utilizing p53 and GADD45 is defective in ataxia-telangiectasia. Cell 71:587–597.

Ketterling RP, Vielhaber E, Sommer SS. 1994. The rates of G:C→T:A and G:C→C:G transversions at CpG dinucleotides in the human factor IX gene. Am J Hum Genet 54:831–835.

Kovach JS, McGovern RM, Cassady JD, Swanson SK, Wold LE, Vogelstein B, Sommer SS. 1991. Direct sequencing from touch preparations of human carcinomas: analysis of p53 mutations in breast carcinomas. J Natl Cancer Inst 83:1004–1009.

Kovach JS, Hartmann A, Blaszyk H, Cunningham J, Schaid D, Sommer SS. 1996. Mutation detection by highly sensitive methods indicates that p53 gene mutations in breast cancer can have important prognostic value. Proc Natl Acad Sci USA 93:1093–1096.

Lane DP. 1992. p53, guardian of the genome. Nature 358:15–16.

Li D, Cao Y, He L, Wang NJ, Gu JR. 1993. Aberrations of p53 gene in human hepatocellular carcinoma from China. Carcinogenesis 14: 169–173.

Li WH, Wu CI, Luo CC. 1984. Nonrandomness of point mutation as reflected in nucleotide substitutions in pseudogenes and its evolutionary implications. J Mol Evol 21:58–71.

Li X, Drost JB, Roberts S, Kasper C, Sommer SS. 2000. Factor IX mutations in South Africans and African Americans are compatible with primarily endogenous influences upon recent germline mutations. Hum Mutat Mutation in Brief Online 367.

Liu JZ, Li X, Drost J, Thorland EC, Liu Q, Lind T, Roberts S, Wang HY, Sommer SS. 2000. The human factor IX gene as germline mutagen test: samples from mainland China have the putatively endogenous pattern of mutation. Hum Mutat 16:31–36.

Livingstone LR, White A, Sprouse J, Livanos E, Jacks T, Tlsty TD. 1992. Altered cell cycle arrest and gene amplification potential accompany loss of wild-type p53. Cell 70:923–935.

Lok E, Ratnayake WM, Scott FW, Mongeau R, Fernie S, Nera EA, Malcolm S, McMullen E, Jee P, Clayson DB. 1992. Effect of varying the type of fat in a semi-purified AIN-76A diet on cellular proliferation in the mammary gland and intestinal crypts in female Swiss Webster mice. Carcinogenesis 13:1735–1741.

Lowe SW, Ruley HE, Jacks T, Housman DE. 1993. p53-dependent apoptosis modulates the cytotoxicity of anticancer agents. Cell 74:957–967.

Manjanatha MG, Shelton SD, Culp SJ, Blankenship LR, Casciano DA. 2000. DNA adduct formation and molecular analysis of in vivo lacI mutations in the mammary tissue of Big Blue® rats treated with 7,12-dimethylbenz[a]anthracene. Carcinogenesis 21:265–273.

Martin FL, Venitt S, Carmichael PL, Crofton-Sleigh C, Stone EM, Cole KJ, Gusterson BA, Grover PL, Phillips DH. 1997. DNA damage in breast epithelial cells: detection by the single-cell gel (comet) assay and induction by human mammary lipid extracts. Carcinogenesis 18:2299–2305.

Martin FL, Pfau W, Cole KJ, Venitt S, Fay LB, Marquardt H, Phillips DH, Grover PL. 1998a. Morphological transformation of C3H/M2 mouse fibroblasts by extracts of human mammary lipid. Biochem Biophys Res Commun 251:182–189.

Martin FL, Pfau W, Venitt S, Carmichael PL, Crofton-Sleigh C, Cole KJ, Stone EM, Marquardt H, Grover PL, Phillips DH. 1998b. DNA strand breaks and cell transformation induced by human mammary lipid extracts. Proc Am Assoc Cancer Res 39:5734–5739.

Martin FL, Cole KJ, Weaver G, Williams JA, Millar BC, Grover PL, Phillips DH. 1999. Genotoxicity of human milk extracts and detection of DNA damage in exfoliated cells recovered from breast milk. Biochem Biophys Res Commun 259:319–326.

Millikan R, DeVoto E, Newman B, Savitz D. 1995. Studying environmental influences and breast cancer risk: suggestions for an integrated population-based approach. Breast Cancer Res Treat 35:79–89.

Morley AA, Turner DR. 1999. The contribution of exogenous and endogenous mutagens to in vivo mutations. Mutat Res 428:11–15.

Petrakis NL, Maack CA, Lee RE, Lyon M. 1980. Mutagenic activity in nipple aspirates of human breast fluid [letter]. Cancer Res 40:188–189.

Prosser J, Thompson AM, Cranston G, Evans HJ. 1990. Evidence that p53 behaves as a tumour suppressor gene in sporadic breast tumours. Oncogene 5:1573–1579.

Saad S, Rowley G, Tagliavacca L, Green PM, Giannelli F, UK Haemophilia Centers. 1994. First report on UK database of haemophilia B mutations and pedigrees. Thromb Haemost 71:563–570.

Saitoh S, Cunningham J, DeVries EMG, McGovern RM, Schroeder JJ, Hartmann A, Blaszyk H, Schaid D, Sommer SS, Kovach JS. 1994. p53 gene mutations in breast cancers in Midwestern U.S. women: null as well as missense-type mutations are associated with poor prognosis. Oncogene 9:2869–2875.

Sands AT, Suraokar MB, Sanchez A, Marth JE, Donehower LA, Bradley A. 1995. p53 deficiency does not affect the accumulation of point mutations in a transgene target. Proc Natl Acad Sci USA 92:8517–8521.

Sarkar G, Yoon H-S, Sommer SS. 1992. Dideoxy fingerprinting (ddF): a rapid and efficient screen for the presence of mutations. Genomics 13:441–443.

Scorsone KA, Zhou YZ, Butel JS, Slagle BL. 1992. p53 mutations cluster at codon 249 in hepatitis B virus-positive hepatocellular carcinomas from China. Cancer Res 52:1635–1638.

Skopek TR, Kort KL, Marino DR, Mittal LV, Umbenhauer DR, Laws GM, Adams SP. 1996. Mutagenic response of the endogenous *hprt* gene and *lacI* transgene in benzo[a]pyrene-treated Big Blue B6C3F1 mice. Environ Mol Mutagen 28:376–384.

Sommer SS. 1990. Mutagen test. Nature 346:22–23.

Sommer SS. 1994. Does cancer kill the individual and save the species? Hum Mutat 3:166–169.

Sommer SS, Ketterling RP. 1993. A postulated mechanism for deletions with inversions. Am J Hum Genet 52:1016–1018.

Sommer SS, Scaringe WA, Hill KA. 2001. Human germline mutation in the factor IX gene. Mutat Res 487:1–17.

Soussi T, Caron de Fromentel C, May P. 1990. Structural aspects of the p53 protein in relation to gene evolution. Oncogene 5:945–952.

Soussi T, Legros Y, Lubin R, Ory K, Schlichtholz B. 1994. Multifactorial analysis of p53 alteration in human cancer: a review. Int J Cancer 57:1–9.

Stemmermann GN. 1991. The pathology of breast cancer in Japanese women compared to other ethnic groups: a review. Breast Cancer Res Treat 18:S67–S72.

Takeshima Y, Seyama T, Bennett WP, Akiyama M, Tokuoka S, Inai K, Mabuchi K, Land CE, Harris CC. 1993. p53 mutations in lung cancers from non-smoking atomic-bomb survivors. Lancet 342: 1520–1521.

Thorland EC, Weinshenker BG, Liu J-Z, Ketterling RP, Vielhaber EL, Kasper CK, Ambriz R, Paredes R, Sommer SS. 1995. Molecular epidemiology of factor IX germline mutations in Mexican Hispanics: pattern of mutation and potential founder effects. Thromb Haemost 74:1416–1422.

Varley JM, Brammar WJ, Lane DP, Swallow JE, Dolan C, Walker RA. 1991. Loss of chromosome 17p13 sequences and mutation of p53 in human breast carcinomas. Oncogene 6:413–421.

Walker VE, Gorelick NJ, Andrews JL, Craft TR, de Boer JG, Glickman BW, Skopek TR. 1996. Frequency and spectrum of ethylnitrosourea-induced mutation at the *hprt* and *lacI* loci in splenic lymphocytes of exposed *lacI* transgenic mice. Cancer Res 56:4654–4661.

Welsch CW. 1994. Interrelationship between dietary lipids and calories and experimental mammary gland tumorigenesis. Cancer 74:1055–1062.

Wink DA, Kasprzak KS, Maragos CM, Elespuru RK, Misra M, Dunams TM, Cebula TA, Koch WH, Andrews AW, Allen JS. 1991. DNA deaminating ability and genotoxicity of nitric oxide and its progenitors. Science 254:1001–1003.

Woutersen RA, Appel MJ, van Garderen-Hoetmer A, Wijnands MV. 1999. Dietary fat and carcinogenesis. Mutat Res 443:111–127.

Wrensch MR, Petrakis NL, Gruenke LD, Miike R, Ernster VL, King EB, Hauck WW, Craig JC, Goodson WH. 1989. Breast fluid cholesterol and cholesterol beta-epoxide concentrations in women with benign breast disease [published erratum appears in Cancer Res 1989 Jul 1;49(13):3710]. Cancer Res 49:2168–2174.

Wynder EL, Cohen LA, Muscat JE, Winters B, Dwyer JT, Blackburn G. 1997. Breast cancer: weighing the evidence for a promoting role of dietary fat [see comments]. J Natl Cancer Inst 89:766–775.

Xue L, Newmark H, Yang K, Lipkin M. 1996. Model of mouse mammary gland hyperproliferation and hyperplasia induced by a Western-style diet. Nutr Cancer 26:281–287.

Yin Y, Tainsky MA, Bischoff FZ, Strong LC, Wahl GM. 1992. Wild-type p53 restores cell cycle control and inhibits gene amplification in cells with mutant p53 alleles. Cell 70:937–948.

Zambetti GP, Levine AJ. 1993. A comparison of the biological activities of wild-type and mutant p53. FASEB J 7:855–865.

Ziegler RG, Hoover RN, Pike MC, Hildesheim A, Nomura AM, West DW, Wu-William AH, Kolonel LN, Horn-Ross PL, Rosenthal JF, Hyer MB. 1993. Migration patterns and breast cancer risk in Asian-American women. J Natl Cancer Inst 85:1819–1827.

Emam Saleh[a], Anke Bockholt[a], Steve S. Sommer[a]

[a]Department of Molecular Genetics, City of Hope National Medical Center, Duarte, CA

SCIMLA, A GENERAL TOOL FOR MEASURING SOMATIC MUTATION LOAD IN SOLID TISSUES: DRAMATIC INTER-INDIVIDUAL VARIATION IN MAMMARY CELLS IS CONSISTENT WITH THE LIPOPHILIC MUTAGEN HYPOTHESIS OF BREAST CANCER

Cancer is a disease of mutations and most if not all mutagens are carcinogens. More rigorous analysis of the relationship between mutations and cancers would benefit from a general method of measuring of mutation load, i.e., mutation frequency, pattern and spectrum, in humans and animals. Single Cell Imunohistochemical Mutation Load Assay (SCIMLA) is a method to measure somatic mutation load from single cells. SCIMLA requires reporter genes in which missense mutations often cause protein accumulation. P53 is the prototype reporter gene; other reporter genes will be developed. For SCIMLA, single cells that stain positively for protein accumulation are microdissected from ethanol fixed and paraffin embedded tissues. The reporter gene is amplified from the single cell and sequenced for somatic mutations. SCIMLA was used to test a first prediction of the lipophilic hypothesis of breast cancer, i.e. mammary cells are uniquely sensitive to a diverse population of lipophilic mutagens in the diet due to the unique anatomy of breast tissue: many tiny islands of cancer-prone mammary cells surrounded by a sea of terminally differentiated fat cells that serve to concentrate lipophilic mutagens. Our observation of marked interindividual variation in mammary cell mutation pattern is consistent with the hypothesis that marked interindividual variation in mutation pattern could reflect dietary exposure to the mutagenic fingerprints of a diversity lipophilic mutagens. Based on these data, a mouse model is being used to examine directly the effect of dietary lipophilic mutagens on mutation frequency and pattern in mammary vs colon and lung epithelium.

Conclusion:

SCIMLA is a tool for making key and unprecedented measurements of somatic mutation from single cells in situ, which may clarify the relationship between mutations and cancer as well as other diseases. The dramatic interindividual variations observed in mammary tissues constitutes the first biological application of this novel technology.